

Engineering organisms for release

Britain's Royal Commission on Environmental Pollution has produced unexpectedly cautious proposals for the regulation of the release of engineered organisms into the environment, but they are none the worse for that.

THE products of genetic engineering have so far been chemicals, materials often indistinguishable from those occurring naturally, but for the past five years there have been plans afoot to use engineered organisms as organisms — as crops, farm animals and the like. For just over a decade, work with genetically engineered organisms in Britain, as mostly elsewhere, has relied on the containment of the organisms. From the outset, in Britain, there has also been a committee (now the Advisory Committee on Genetic Manipulation) which has considered all proposals to produce genetically engineered organisms, mostly microorganisms. Now it is proposed that there should be a new committee, empowered by new legislation, to license all plans to release engineered organisms into the environment (see page 84).

The present arrangements for dealing with engineered organisms in the laboratory and in industrial plants have several virtues. One is their adaptability. As experience has accumulated, it has been possible to relax the precautions judged necessary for some (but not all) manipulations so as to protect the health of people (chiefly those working in the laboratories concerned). By arranging that the central committee is mirrored by local committees, often at the laboratory level, these arrangements have also helped to spread a general sense of confidence in the good sense of such genetic manipulation as is allowed. The Royal Commission plainly seeks to capture as many as possible of the social benefits of the old arrangements for the novel circumstances in which organisms are deliberately released, so that containment is, by definition, no security. That is a proper starting-point.

The difficulties will arise only later. The commission's case for new legislation is simple (and undeniable): present legislation (in Britain) covers people, not the environment. Specifically, the commission asks that there should be a statutory licensing procedure (rather than the present process of less formal consent) for each successive stage in the development of an engineered organism intended for release. There would also be, in the commission's scheme, an "imaginative" procedure for monitoring the environmental effects of releasing organisms into the wild. The commission also asks that information about the intentions of researchers and commercial producers should be made publicly available — and that patent legislation should, if necessary, be amended so as not to prejudice designers' rights. The good sense of these

proposals will hang crucially on the smooth functioning of the mechanism to which they lead. That is an issue to which the commission should have paid more attention.

The case for licensing as such is strong; there is a need to know what field trials have been carried out. But it is also necessary that the licensing procedure should be as informal as particular circumstances require, and flexible as well. When, for example, half a dozen differently engineered *Rhizobium* bacteria have been found to increase nitrogen fixation, but only temporarily, and when some understanding of the reasons has been arrived at, the granting of licences for similar releases should become easier and monitoring less onerous. Will that be possible when there has been legislation specifically on this subject? The obvious difficulty for future governments is that there are no votes in standing out against allegations, however unfounded, that even a familiar engineered organism may play havoc with a novel environment. It will take clear-headed drafting of the new legislation, and self-denying delegation of responsibility to independent committees, if the British government is not to stumble into a legislative regime that restricts the potential as well as the dangers of the new technology.

A further difficulty hangs on what is meant by protection of the environment. The introduction to the commission's report properly draws attention to the environmental changes brought about by the introduction of exotic plants and animals into Britain. Dutch elm disease is fresh in most people's minds, and has plainly been a disaster. But must the same be said of the rhododendron species cherished by Victorian gardeners but now running wild through many natural landscapes? In a largely artificial and intensively managed landscape such as that, taste may determine whether environmental changes are beneficial or the opposite. The danger with genetic manipulation is that change will become synonymous with catastrophe.

The Royal Commission, like many other British institutions these days, is also reasonably at odds with the European Community's proposals. The position is that the Community is still considering the European Commission's draft directives on the release of organisms, but is likely to require (in this field as in others) that genetically engineered organisms approved for sale in one of the member states should be allowed for sale in all the others (subject to derogation on geographical grounds). The



European Commission's point is that the environment varies widely in Europe, and that organisms whose release is considered safe in one state may be harmful in another. The proposed safeguard is that all member states will have the right to protest when experimental releases anywhere are planned, but the Royal Commission evidently believes that this will not suffice. The same problem will of course arise elsewhere, which is why there is a case for giving European institutions a stronger role in the licensing procedure than that now planned. □

Can heresy be real?

INSERM has done the only decent thing by confirming Dr Jacques Benveniste in his post. But what happens now?

THESE have been wretched months for Dr Jaques Benveniste, the director of INSERM's unit for the study of immunology and allergy, otherwise called INSERM 200. During the past few months, as part of the research council's routine quadriennial review of its establishments, (see page 89) two separate committees have trudged through his laboratory, each of them mixing scepticism with admiration in their reports. The admiration is natural enough. Since 1981, when Benveniste became the unit's director, his laboratory has produced a substantial amount of solid research, which has been published in reputable journals. But, sadly, Benveniste is best known for what seems to be an aberration — the assertion that some biological reagents retain their activity even when indefinitely diluted, well beyond the point at which all molecules should have disappeared. Outwardly, these two roles conflict. INSERM deserves credit for recognizing that, in science, either the concept of heresy is meaningless or, otherwise, heresy cannot be an excommunicable offence.

The Benveniste business is complicated, as readers of this journal know only too well. A year ago, *Nature* published a paper (E. Davenas *et al.* **330**, 816; 1988) from Benveniste's group, following it with an almost equally contentious document, based on a visit to INSERM 200, which was critical of the procedures by which data had been gathered and interpreted (*Nature* **334**, 287; 1988). Since then, Benveniste says, the statistical treatment of the high-dilution data has been improved — and the phenomenon persists. Benveniste has nevertheless acceded to an earlier demand by INSERM that his high-dilution experiments should not consume public resources. He now also agrees to refrain from conducting science through the daily press and to keep an "open mind" to alternative interpretations of his high-dilution data. He remains convinced that his observations point to a real phenomenon, and protests that freedom in science is jeopardized by the pressures to which he has been subjected.

Nobody will dispute people's right to be wrong. Mistakes happen all the time. It is also proper that people

should persist with their beliefs in the face of scepticism. Where would the heliocentric hypothesis be had Galileo not done so? There is also an honourable tradition that laboratory directors, often appointed because of their achievements in some field, should provide their colleagues with encouragement and resources while themselves embarking on more adventurous if risky projects.

Benveniste's essential difficulty is that, even after the unusual attention paid to his claims, few seem openly to have followed his line of investigation. (*Nature*, in the past year, has been sent for publication a single paper reporting similar observations with a different system, now with its authors for further clarification, but may well have discouraged others by its treatment of Benveniste's contribution.) Even the three collaborating laboratories quoted by Benveniste a year ago have been publicly silent. Although truth in science cannot be determined by casting votes, Benveniste's failure to recruit open support within the scientific community should be discouraging.

INSERM, an excellent organization with no inclination to be classed with the Spanish inquisition in the suppression of heterodoxy, has resolved its dilemma excellently. Considerations other than freedom in research determine the relationship between a research organization and its laboratory directors. Benveniste cannot be faulted as an administrator and motivator — his laboratory last year was as enthusiastic for its orthodox as for its unorthodox work. Nor should it be a consideration that a single director holding views widely disbelieved could damage the reputation of French science (as some have darkly suggested): after a decade of remarkable progress within France and internationally, only narrow chauvinists could believe that. But it does make sense that Benveniste should now be asked, with breathtaking openness, not to rock INSERM's boat uncomfortably by over-zealous and even surreptitious advocacy of his cause.

The truth about heresy in science, after all, is that nobody can prove that heretical beliefs are in any sense wrong. The standard complaint against the popperian view of science — that there is no such thing as truth — can also be turned inside out, to show that, in the fields in which hypotheses cannot yet be tested, there is no such thing as being wrong. That, in a court of philosophical law, would be Benveniste's best defence. In the real world, his best defence is simpler: people who believe that water can have a memory and things like that, but who are otherwise reputable researchers, have a right to expect that their colleagues will listen to them courteously. Correspondingly, they also have a duty to listen when their critics say that, if that is the hypothesis, this is how it can be tested convincingly. But Benveniste and his critics have usually dealt with each other by pretending that the others do not exist. Should they not now, in the interests of good manners if not of truth, break with precedent? Benveniste is wrong about the memory of water, but has won the right to expect that people will put him straight. □

Abortion ruling divides the United States

- Attempts to redefine beginning of life
- Opposing groups ready for battle

Washington

LAST week's US Supreme Court ruling on the constitutionality of a Missouri law that places new restrictions on the right to an abortion appears to have sundered the connection between the legal view of abortion and the medical view of the course of pregnancy. In so doing, it opens the door to a host of further legal cases that may both restrict further the availability of abortion and challenge biomedical practices in many other areas, from the use of *in vitro* fertilization to the treatment of fetal abnormalities.

The ruling, delivered last Monday by a majority of five Supreme Court justices to four, has only a small direct effect, requiring Missouri doctors to test whether a fetus thought to be 20 weeks or older has reached a stage of development at which it could live outside the womb before considering an abortion. That ruling is wholly uncontentious. But within the opinions written by the Supreme Court justices are a host of indications that they do not accept the logic of the *Roe v Wade* decision that has determined US abortion rights for the past 16 years. Most significantly, the majority of justices let stand a preamble to the Missouri act that sets out that life begins at conception, and backed an opinion that could extend state interest to the whole period of a pregnancy.

Before last week's ruling, the connection between the state's right to regulate abortion and the course of pregnancy appeared solidly based in biological commonsense. In 1973, *Roe v Wade* accepted, as Chief Justice John Paul Stevens puts it, that "there is an obvious difference between the state interest in protecting the freshly fertilized egg and the state interest in protecting a nine-month gestated fully sentient fetus on the eve of birth. There can be no interest in protecting the newly fertilized egg from physical pain or mental anguish because the capacity for such suffering does not yet exist; respecting a developed fetus, however, that interest is valid".

In essence, *Roe v Wade* saw an increase in moral objections to abortion as a fetus developed, with a landmark at about 24 weeks when a fetus becomes capable of survival outside the womb. *Roe v Wade* divided pregnancy into three trimesters. In the first, women were given a right to abortion, protected by the constitutional right to privacy. In the second, the state was allowed some say in the control of abortion, but only in such regulations that

protected the woman, not the fetus. In the final trimester, as the fetus became viable outside the womb, the state could regulate and even proscribe abortion.

On the surface, the new ruling does not change this state of affairs, the majority of justices claiming it simply "modifies and narrows" the *Roe* ruling. But in a single short paragraph, Chief Justice William Rehnquist opens the door to an onslaught on the logic of *Roe* by saying "we do not see why the state's interest in protecting human life should come into existence only at the point of viability and that there should therefore be a rigid line allowing state regulation after viability but prohibiting it before viability". Consistent with that view, the majority let stand the preamble to the Missouri act which declares that "the life of each human being begins at conception" and that "unborn children have . . . all the rights, privileges and immunities available to other persons, citizens and residents of this state".

The Missouri act defines conception as "the fertilization of the ovum of a female by the sperm of a male", even though implantation does not occur until at least six days after fertilization. It therefore implies regulation not only of abortions in the first two trimesters, but also of forms of contraception such as the IUD and the morning-after pill. It would also appear to grant rights to all eggs fertilized in a petri dish for *in vitro* fertilization, and makes impossible the disposal of excess embryos.

The preamble passed the scrutiny of the majority of justices on the grounds that "it will be time enough for Federal courts to address the meaning of the preamble should it be applied . . . until then, this Court is not empowered to decide . . . abstract propositions". Only Justice John Paul Stevens, who dissented from the majority view, saw the preamble as essentially a theological argument that was invalid under the First Amendment and was "an unequivocal endorsement of a religious tenet of some but by no means all Christian faiths".

By striking at the foundations of *Roe v Wade*, but offering no clear guidelines in its place, the supreme court invites an avalanche of state legislation that will attempt to test further the right to control abortion. Behind that legislation will be vigorous pressure groups, many with roots in the evangelical Christian movement, that believe abortion is permissible only to save the life of the mother, and not even in cases of rape or incest.

Opposing them will be a host of women's groups. Justice Harry A. Blackmun, who dissented from the ruling, spoke for them when he wrote it would "clear the way once again for government to force upon women the physical labour and specific and direct medical and psychological harms that may accompany carrying a fetus to term . . . every year, many women, especially poor and minority women, would die or suffer debilitating physical trauma, all in the name of enforced morality or religious dictates or lack of compassion, as it may be".

The pro-abortion lobby believes that a majority of women will come forward to ensure that the freedoms they now enjoy will not be taken away from them. But the anti-abortion lobby believe they alone have the necessary staying power. As *Lifeletter*, the newsletter of one such organization bluntly put it, "pro-aborts have been notoriously unable to mobilize the kind of in-depth grass-roots, long-term tenacity that has sustained anti-aborts . . . Jane Fonda and Co. have a limited attention span . . .".

Alun Anderson

NUCLEAR POWER

Plant goes for a dollar

Boston

SHAREHOLDERS of the Long Island Light Company (LILCO) have voted to close the completed, but never operated, Shoreham nuclear power plant. It is the first time a utility has ever abandoned a licensed plant before it has gone on-line.

The shareholder vote gave almost unanimous support for a shutdown plan which allows the utility to divest itself completely of responsibility for the plant by selling it to the state of New York for \$1. Under the agreement, the state would then be responsible for decommissioning the plant.

But Shoreham's fate may still not be sealed. After the announcement of the LILCO shareholder's decision, US Energy Department deputy secretary W. Henson Moore said that the vote will not affect his department's efforts to keep the plant alive. Energy Secretary James D. Watkins has on several occasions decried the closure of Shoreham and promised to make its operation a personal priority.

According to an Energy Department spokesperson, the agency will seek to block the transfer of Shoreham to New York state during Nuclear Regulatory Commission hearings later this year. But that move may come too late. LILCO officials have already announced that they will begin immediately to withdraw the plant's uranium fuel rods, a process that should be completed this summer.

LILCO vice president Joseph W. McDonnell expresses relief at the latest development: as far as the utility is concerned, he says, the "controversy has gone on much too long."

Seth Shulman

Controls needed on release

London

NEW legislation to back up a system of compulsory registration and licensing for any release in Britain of a genetically engineered organism is advocated in a report published last week by the Royal Commission on Environmental Pollution (RCEP). To release an organism without a licence should be a criminal offence.

The commission's long delayed but carefully considered report argues against immediately categorizing any types of release as sufficiently free of risk not to require individual scrutiny by experts and licensing by the Secretary of State for the Environment and the Health and Safety Commission (HSC). But it recommends against any moratorium on releases.

In launching the report, Lord Lewis, RCEP's chairman, said that members of the commission were persuaded that too little is known of the possible risks of releases to proceed other than in a very cautious manner. Each organism "will probably pass through several stages of experimental development and trial release", says the report, and "each stage . . . should be the subject of a licence". A further licence would be necessary for any product containing the organism.

Pointing out that the biggest brake on the acceleration of releases, which now number about 80 worldwide, would be a case of serious damage caused by an inadequately scrutinized experiment, the commission says that "some may consider our proposals onerous but we believe them to be necessary". Should the strict controls it proposes drive people to use methods other than genetic engineering to modify organisms, particularly micro-organisms, these too may have to be controlled by legislation, the report suggests.

The commission became convinced that environmental monitoring of all releases is of paramount importance, said Lord Lewis. The report says that releasers must be required not only to monitor and report their results for the agreed duration of the experiment, but to continue monitoring for "an appropriate period" thereafter, with a degree of imagination to catch any unexpected signs of damage. Compliance would be checked by appropriately trained inspectors, while Britain's army of amateur natural historians could be expected to notice any untoward effects.

Members of the commission had nothing but praise for the committee of experts that has already overseen the

handful of releases in Britain in the three-year gestation period of the report. They recommend that this committee should be constituted in its own right to advise the Secretary of State for the Environment and the HSC on each proposed release. The committee should also produce codes of practice, advise on the need for research and the possibility of categorizing releases, and produce an annual report.

Public access to information on releases is essential, says the commission. There should be a public register of applications for experiment and product licences, and public advertisements of proposed releases. The public should also have access to the information on the basis of which the expert committee has made a recommendation, taking into account any need for commercial confidentiality.

In immediate reaction to the RCEP report, both Professor David Bishop, who has completed four UK releases of engineered baculoviruses, and UK Genetics Forum, a new group that aims to focus public concern over biotechnology, criticized the lack of public interest group representation on the expert committee proposed. The UK Genetics Forum is also disappointed that the commission did not support its call for a partial moratorium on releases, banning those with commercial aims while allowing those designed to assess risks to proceed.

Should RCEP's recommendations become the basis of new legislation, Britain would have a law largely similar to that in Denmark, the only existing European law, said commission member Lord Nathan. (The first Danish releases have just been approved — see *Nature* 339, 653; 29 June 1989.) The British law would, however, be considerably more restrictive than that proposed by the European Commission, and now the subject of considerable debate in the European Parliament and Council (see *Nature* 339, 653; 29 June 1989). RCEP's report criticizes the European proposal for its "very extensive list of exclusions" in respect of the products that would require control and for its 'single market' proposals that a product approved in one European Community country should not be restricted in others. This, the report argues, ignores the very different geographical environments in Europe.

RCEP's recommendations fall within the possibilities for new legislation very recently put forward in a Department of the Environment consultation paper (see *Nature* 339, 499; 15 June 1989) but are at the stricter end of the spectrum. The department will be considering reactions to both documents after the summer and hopes to finalize new legislation to control the release of genetically engineered organisms in time to include it in a general environment bill that may be ready later this year.

Peter Newmark

Experimental releases in progress in Britain

London

THE Royal Commission on Environmental Pollution had hoped to finish its report (see above) within a year or so but took three years. Commission member and economist Professor Aubrey Silberston said this was because the topic was unexpectedly complex and he and his fellow members were largely ignorant of it before they began. (Of the commission's 14 members, only Professor William Stewart has been an experimental biologist.) While the members were acquiring their knowledge, researchers were applying theirs to the design and execution of several release experiments.

The first release in Britain, in 1986, was a baculovirus engineered to contain an innocuous genetic marker. This initiated a series of releases aimed eventually to increase the efficiency of baculoviruses as biological control agents. The work is being carried out by Professor David Bishop and colleagues at the Natural Environment Research Council's Institute of Virology and Environmental Microbiology in Oxford with help from the Department of the Environment (DoE) and the Biotechnology Action Programme of the European Commission. In the latest release, the virus contains a bacterial 'reporter' gene in place of its coat-protein gene.

The action programme also supports risk-assessment research on plant-surface bacteria at the Oxford institute and on

rhizobial bacteria at the Agricultural and Food Research Council (AFRC)'s Institute of Arable Crops Research. And DoE will this year spend a total of £0.67 million on similar research at several universities and at the Freshwater Biological Association. It hopes to double its budget next year.

A new source of finance for risk-assessment experiments is the Planned Release of Selected and Modified Organisms (PROSAMO) programme, which is half paid for by a consortium of companies and half by AFRC and the Department of Trade. A three-year £1.3-million budget is divided between microbiological research at the Universities of Aberdeen and Essex, and plant research at AFRC's Institute for Plant Science Research and Imperial College, Silwood Park. The plant research involves the release of engineered potato plants, testing in particular the extent to which a bacterial gene engineered into potatoes can be transferred to related species by pollination.

The Royal Commission itself, with support from DOE, has supported some theoretical research to explore whether the successful HAZOP system of exposing unexpected hazards in the operation of chemical plant can be adapted to genetic engineering. GENHAZ, as it is called, has not yet been satisfactorily worked out but may be tested in a practical exercise sometime next year.

Peter Newmark

GENETIC ENGINEERING

West Germany eases law

Munich

SCIENCE organizations breathed a sigh of relief last week in West Germany when the government indicated that it would ease some of the most restrictive provisions in a new 'basic law' to regulate genetic engineering. But the rest of the law may still be harsh enough to accelerate the exodus of pharmaceuticals research and development from the country.

The most welcome change for science organizations, including the Deutsche Forschungsgemeinschaft (DFG) and the Max Planck Society (MPS), was the dropping of a requirement for licensing of individual experiments with genetically modified organisms classified at 'Level 1', or 'not dangerous'. DFG and MPS had warned at a closed hearing in Bonn on 25 May that such a licensing procedure could become a bureaucratic nightmare (see *Nature* 339, 327; 1 June 1989).

The regulations for handling riskier organisms or strains were also relaxed, making a general licensing procedure possible. A central licensing commission (ZKBS) will be called upon to assess experiments within 60 days after an application is submitted.

Representatives of DFG and MPS familiar with the general outlines of the new law continue to object to the regulation of genetic engineering as if it were a dangerous technology. DFG has offered to help the government to draft the all-important regulations that will be used in applying the law in specific cases.

The effect of the new law on industry will depend largely on the amount of public participation allowed in licensing procedures. In the current version, the licensing of production facilities that use

hazardous organisms will require public participation, which could introduce years of delay into the licensing procedure. Drawn-out licensing procedures are one of the main justifications for pharmaceutical producers such as Hoechst and BASF pulling some research and development out of West Germany.

A struggle is brewing in Parliament about this question. The cabinet is expected to submit the law to Parliament on 12 July. The *Länder* (states), which are represented in the Bundesrat (upper house), are expected to push for more regulatory power. Industry supports a more central administration, with regulatory power concentrated in the ZKBS or the Health Ministry. Allowing the *Länder* to make their own rules regarding public participation would be "fatal", said a spokesman for the chemical industry consortium VCI.

An agreement reached in the European Communities (EC)'s council of environment ministers on 8 June provided inspiration for the changes in the West German draft law. The new EC regulations (which have yet to take effect) will set minimum standards for genetic engineering research in the laboratory; standards for the release of genetically modified organisms into the environment still remain to be decided (see *Nature* 339, 653; 29 June 1989).

There is still a climate of strong distrust between researchers and environmentalists in West Germany, further clouding the future of molecular biology. In a recent parliamentary debate, a Green Party member demanded that all genetic engineering be banned because it could be "even more of a threat to mankind than nuclear energy".

Steven Dickman

CONSERVATION

Japanese traditions keep ivory trade alive

Tokyo

JAPAN, the world's largest importer of ivory, has instituted a partial ban on ivory imports which should significantly reduce the nation's consumption and help international efforts to protect dwindling populations of elephants. But Japan seems unlikely to follow European nations and the United States in imposing a total ban on imports (see *Nature* 339, 494; 1989).

Japan's Ministry of International Trade and Industry (MITI) announced last month that imports will be allowed only direct from the country of origin. As nearly 90 per cent of Japan's ivory imports last year came indirectly from suppliers in places such as Hong Kong and Singapore, many of whom are believed to have acquired the ivory illegally, the partial import ban should help to stop ivory poaching and smuggling.

Japan imports almost half of the world's

ivory. Most of it is used to make personal seals (*hanko*) and plectrums for shamisen — a traditional three-string musical instrument played by geisha. The Japan General Merchandise Importers Association bemoans MITI's decision, saying it will harm Japan's traditional culture. And Japan's ivory agents sent delegates to a meeting of the Convention on the International Trade in Endangered Species (CITES) in Botswana last week to plead their case.

The Botswana meeting was held to discuss the possibility of instituting a total ban on all trade of African elephant ivory by upgrading the African elephant from Appendix 2 of CITES, which allows limited trade, to Appendix 1.

Japan is also expected to oppose a total ban when the 102 parties of CITES meet in Switzerland in October to vote on the issue.

David Swinbanks

ANIMAL PATENTS

Oncomouse seeks European protection

Munich

EUROPE moved one step closer to a showdown on the question of whether animals may be patented when the European Patent Office (EPO) in Munich on 28 June rejected an application to patent a transgenic mouse, known as 'oncomouse', that carries human cancer genes, or oncogenes.

But the real battle will be fought in the coming months, after European Community (EC) member states have responded to a controversial proposal from the EC executive body, the European Commission. The proposal, developed in the economics directorate of the Commission, calls for more patent protection of biotechnological discoveries, including animals and plants altered through genetic engineering. The EC Council of Ministers has begun expert hearings on the subject. West Germany and Denmark are expected to oppose the measure, and Britain will probably support it. Agreement is "not yet in sight", said Dirk Brouer of the West German Justice Ministry.

Meanwhile, the agriculture directorate of the Commission is preparing a second proposal for a new "species protection law" in the EC. The Commission has decided to wait until this proposal has been completed — it has been promised for December 1989 — before redrafting the first proposal or submitting it to the European Parliament.

Currently, the European Patent Convention, signed by 13 European countries including most EC members, forbids the patenting of "plant and animal varieties".

Oncomouse was developed by Harvard professor Philip Leder and Timothy Stewart, now of Genentech, and patented in April 1988 in the United States by Harvard University (see *Nature* 332, 668; 1988). The patent attorney for Harvard, Paul Clark, said that there will be an appeal based on the definition of "varieties", which he claims does not rule out the patenting of transgenic animals such as the so-called oncomouse.

The decision came as no surprise to Harvard and the chemical company Du Pont, which holds exclusive rights to the US patent. In February, EPO had indicated that the application would be rejected and that EPO did not consider itself competent to decide such a crucial issue.

The case is seen as an important test for the European biotechnology industry. Should the EPO ruling eventually be overturned, it could open the door for the patenting of any species — from wheat to pigs — that has been genetically engineered. Environmentalists have defended the status quo, fearing that the granting of patents could concentrate too much power in the hands of multinational companies, to the detriment of the farmer and the developing world.

Steven Dickman

New moon for Neptune

Washington

Six weeks away from its rendezvous with Neptune, the Voyager 2 spacecraft has discovered a third moon orbiting the planet. Neptune's other two moons are the almost Earth-sized Triton and the smaller Nereid. The new object was first seen three weeks ago as a bright spot moving against the background of stars, but no resolved image of it has yet been made. It is estimated to be between 125 and 400 miles in diameter, and is moving in an approximately circular, 60,000-mile-radius equatorial orbit. D.L.

From moulds to money

London

A SLIME-MOULD biochemist is to become the next director of the London School of Economics (LSE). But it is presumably for his other skills that Professor John Ashworth has landed the job. Currently vice-chancellor of the University of



round by attracting an influx of industrial money, Ashworth served as chief scientist in the Central Policy Review Staff, the prime minister's 'think-tank', from 1976 to 1981. He succeeds Dr Indraprasad Patel at LSE in October 1990. P.N.

Embezzlement arrest

London

AN employee of the Chinese Academy of Sciences is under arrest in the biggest embezzlement case in the history of the Chinese People's Republic, it was announced last week. Shen Xiaoping, described as a purchasing agent of the academy's Microelectronics Centre, is alleged to have embezzled more than 390,000 yuan (£115,000) of public money allocated to the centre between March 1985 and December 1987. In all, 12 people are involved in the case. Six are in custody, and the rest are on bail awaiting trial. V.R.

The end of a scientific era?

Munich

THE glowing fifteen-year tradition of the Dahlem Conferences in West Berlin may be brought to an abrupt stop next year unless a new sponsor can be found. The current sponsor, the industry group "Donors Association for Promoting Arts and Sciences in West Germany" (*Stifterverband für die Deutsche Wissenschaft* or SV) plans to keep its longstanding promise to drop the conferences at the end of 1989.

The Dahlem conferences have been praised by participants for their unique structure. There are four conferences a year for 40 to 50 participants, by invitation only, on subjects ranging from neurosciences to global change. By contrast with the slide-projector and coffee-break routine of most scientific meetings, participants meet in small groups to discuss interdisciplinary topics. The result of each conference is a book, written while the discussion is still fresh.

Members of the scientific board of the conferences reacted with outrage at the revelation late last month that the staff of seven had been dismissed from December 1989. Jean-Pierre Changeux of the Pasteur Institute in France said it would be a "disaster" if the conferences were to end; Hubert Markl, president of the West German granting agency DFG (*Deutsche Forschungsgemeinschaft*) called the conferences "an invaluable instrument of international science promotion". DFG provides money for travel grants to the participants in one Dahlem conference a year.

SV had announced in 1986 its intention to transfer control of the conferences to some other agency. At first, the Wissenschaftskolleg (Institute for Advanced Studies) of West Berlin, the construction of which SV helped to finance, seemed prepared to take the reins, but talks broke down in 1986.

The ill-fated Academy of Science and Technology in Berlin offered in 1987 to take over the conferences, but conference leader Silke Bernhard balked at the transfer. The academy plan had to be shelved in April this year when the newly elected Green-Social Democratic government in West Berlin announced its intention to shut down the Academy.

Bernhard, a physician with a strong background in science promotion, is given credit by many participants for the success of the meetings. She saw her autonomy, and with it the quality of the conferences, threatened by the intentions of the Academy. She was supported in her stand by other members of the scientific board, including Changeux and Sir Gordon Wolstenholme of the Royal College of Physicians in London, who have written

stirring letters in her defence.

Bernhard charges that SV has decided to disband the conference organization now, instead of seeking a new sponsor, out of a desire for personal revenge against her and possibly for political reasons. There had been an earlier dispute, concerning publication of the Dahlem conference books, which led to bad blood between Bernhard and the SV general secretary, Horst Niemeyer, and the conference administrator, Hans-Henning Pistor.

Bernhard sees a political motive in the desire for SV to retaliate against the new West Berlin Senate for shutting down the Academy of Science and Technology. Although SV has no official political affiliation, it is thought to be in alignment with the conservative West German Christian Democratic party, which had shepherded the Berlin Academy into existence over the objections of the Social Democrats.

Pistor rejected vehemently Bernhard's charges of vengeance, saying it was beneath the dignity of SV to respond. He said that Bernhard is "an employee and she just has to accept that role. It was solely her refusal to accept the Academy's conditions" that led SV to carry out its plan to end the conferences.

As for the political motive, Pistor called the idea "foolish", saying that SV would "never even think of such an idea". Pistor, who has administered the conferences for SV since 1978, said that "no-one plans" to shut down the conferences if it can be helped. The West Berlin Senate has proposed that the Free University of Berlin take over sponsorship, an idea that Pistor called "difficult" but which the university is still considering.

Another idea which might save the conferences is the setting up of an independent, registered organization, or *Verein*, to administer the conferences and take all legal responsibility. Pistor said that SV would not object to this proposed structure if appropriate board members could be found.

Ironically, money is not the largest problem in keeping the conferences running. The city of West Berlin has promised DM 1 million (about \$500,000) a year for five years toward operating costs, as long as a suitable sponsor is found. DFG, the multinational company Unilever and other sources provide most of the rest of the budget, leaving a shortfall of just DM 160,000.

But time is running out for the conferences. Bernhard expects most of the staff to leave before the end of the year, even though two more conferences are scheduled in November and December.

Steven Dickman & Jennifer Altman

ASTRONOMY

New telescope agreed

Washington

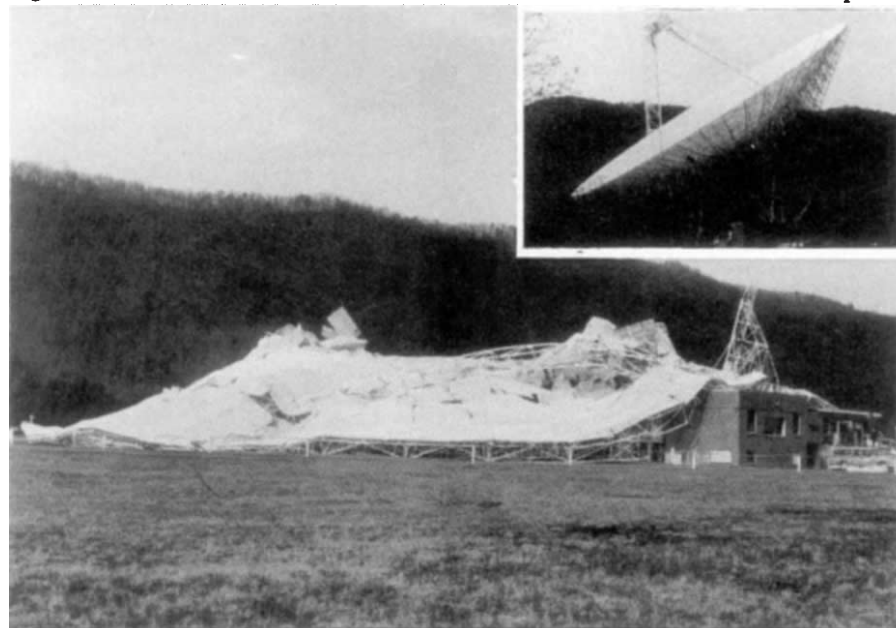
THANKS to the legislative legerdemain of West Virginia Senator Robert C. Byrd, the radiotelescope at Greenbank that collapsed unexpectedly last November (see *Nature* 336, 296; 1988) is to be replaced with a new and more sophisticated design.

Last week, Byrd succeeded in adding a \$75-million budget appropriation for a new telescope to a 1989 Dire Emergency Supplemental Bill that had to be rushed through Congress to provide \$1,200 million in emergency funding for the Veterans' Administration. Now that the bill has been signed into law, the National Science Foundation (NSF) receives an immediate bonus of \$37.5 million on top of its normal budget. The remaining \$37.5 million will be available when the new fiscal year begins on 1 October.

and provide high performance at 7-mm wavelength, with some performance down to 2.5-mm wavelength.

Details have not yet been decided and AUI may go for an unconventional design which would have the feeder, which sends the reflected signal from the main mirror to the electronic detectors, set off to one side. This would give an unblocked main aperture, and would also reduce the size of the sidelobes in the receiver's pattern of sensitivity, a particularly useful attribute in mapping large-scale hydrogen emission from the Galaxy.

A preliminary design will be completed this year, with engineering drawings ready next year. The telescope should be ready by 1995. It will be especially good at looking for millisecond pulsars because, unlike the non-steerable telescope at



Greenbank: the 300-foot radiotelescope before (inset) and after its collapse last November.

Replacing Greenbank was not the NSF's own top priority as it believed there were other, better ways to spend \$75 million than on an updated version of the 26-year-old, 300-foot telescope (see *Nature* 338, 448; 1989). But, as the only major scientific facility in his home state, Greenbank was at the top of Byrd's own list. Using his powerful position as chairman of the Senate Appropriations Committee, Byrd was able to please everyone by ensuring that the telescope would be rebuilt with funds that did not come out of the annual NSF budget.

Paul A. Vanden Bout, director of the National Radio Astronomy Observatory at Greenbank, says that a proposal for a new 100-metre telescope has already been submitted to NSF by Associated Universities Inc. (AUI), the consortium that runs the observatory. Unlike its predecessor, the new telescope will be fully steerable

Arecibo (Puerto Rico), it will be able see towards the Galactic Centre, where globular clusters containing millisecond pulsars are concentrated. The telescope's high performance at short wavelengths should also produce some firsts in the study of molecular clouds in distant galaxies, by allowing the detection of carbon-monoxide emission lines at unprecedentedly high redshifts.

In the future, the telescope could be linked to satellites that Japan and the Soviet Union plan to launch by 1995, making it the terrestrial end of a space-based very-long baseline array.

The budget success for Greenbank effectively ends alternative plans for keeping the facility alive by placing the eastern end of a proposed Laser Interferometry Gravity Wave Observatory (LIGO) at the site. Greenbank is not the most preferred location.

Alun Anderson

BIONET

NIH to end support for electronic network

Washington

THE US National Institutes of Health (NIH) intends to pull the plug of the BIONET molecular biology electronic network at the end of the summer. In a message flashed across the network's electronic bulletin board last week, IntelliGenetics, the California biocomputing company that runs BIONET, advised its users at 925 laboratories to make "alternative" arrangements because NIH support will end on 30 September, forcing the company to shut down the network or restrict its services.

IntelliGenetics set up BIONET five years ago under a \$5 million "cooperative agreement" with NIH. Since then, the network has provided electronic mail services, access to DNA and amino-acid sequence and protein coordinate databases, and sequence-manipulation software to more than 2,000 users for a modest fee of \$400 per year.

But BIONET's money ran out earlier this year, and the NIH advisory committee that oversees the network declined to continue funding on the grounds that its research programme was weak. Unlike a contract, where a service is performed for a set fee, an NIH cooperative agreement requires grantees to conduct research programmes of the same calibre as those financed under competitive grants.

BIONET users have reacted strongly to the announcement. David Kristofferson, manager of BIONET at IntelliGenetics, has already received 25 electronic mail messages objecting to NIH's decision, and says his phone has been ringing constantly. An open letter soliciting comments on NIH's decision posted on the network by Dan Davison, a staff member from the GenBank DNA sequence database at Los Alamos National Laboratory, has collected nearly 50 responses from around the world.

But BIONET users may not be left totally out in the cold: Charles Coulter, head of the NIH office that administers BIONET, says he expects the "minimum capabilities" of BIONET to be transferred to the GenBank database. IntelliGenetics also runs GenBank, under an NIH contract, with support from the Department of Energy. NIH will decide next month whether to expand GenBank to include some of BIONET's services.

Kristofferson says beefing up GenBank may satisfy the needs of some users, but he fears that small laboratories may suffer. Now, anyone with a desktop computer and a modem can tap into BIONET's mainframe computer to sort sequences in GenBank without paying for a long-distance telephone call. Without BIONET, he says, "one way or another, people are going to pay more money".

Carol Ezzell

New regulations proposed

Washington

New federal proposals to regulate the care and use of animals in laboratory experiments have outraged the US biomedical research community, prompting more than 5,000 written complaints to the Department of Agriculture Animal and Plant Health Inspectorate Service (APHIS).

The proposed regulations are described as useless, time consuming and expensive. APHIS is accused of misinterpreting the statute from which the regulations stem and is urged to scrap its proposals and start again. But animal-rights groups, which galvanized 2,000 responses from members of the public, complain that the regulations are not tight enough.

By contrast with the existing animal-welfare policies of the Public Health Service (PHS) and the National Institutes of Health (NIH), APHIS proposes minimum required standards for the design of laboratories and experiments involving animals. These include regulation of the temperature, humidity and lighting of enclosures; a requirement that dogs must be able to see and hear other dogs and must be released once a day for exercise for a period of 30 minutes; and regulations to improve the "psychological well-being" of non-human primates.

Researchers agree that new animal-welfare regulations are needed to ensure compliance with existing guidelines but are opposed to the imposition of rigid and arbitrary design standards by an outside agency. Under the existing policies, institutions determine how the recommended standards are met. By proposing regulations which conflict with the approach of the NIH and the PHS, the Department of Agriculture is accused of exerting more authority than it was granted by Congress under the 1985 Animal Welfare Act amendments. John Miller, of the NIH Office for the Protection from Research Risks, says the NIH is trying to convince APHIS to adopt an approach in line with their own. They have had some success and the policies agreed so far are similar to existing ones, he said.

The Humane Society of the United States welcomes the proposed regulations. Martin Stephens, director of the group's Laboratory Animals Department, says that "in an ideal world, performance standards would be better than design standards, but it would be very difficult to ensure adherence to them". But he argues that there are too many loopholes in the regulations and says that some may have little impact on animal welfare. The exclusion of birds, mice and rats from the regulation would cause "a serious outcry" from the public, he said. APHIS says it is considering regulations to include them.

Under the new regulations, most research institutes would have to renovate or replace cages and equipment and construct new housing facilities to meet increased space requirements, but no money has been appropriated by Congress to cover the costs. APHIS estimates the necessary capital expenditure to be almost \$900 million and maintains that since current spending on biomedical research is more than \$12.8 billion per year, the community can afford the increased costs.

But the Pharmaceutical Manufacturers Association (PMA) says that for its members alone the new regulations would cost \$345 million and that the APHIS estimate is several times too low. The PMA also complains that the regulations would "entangle researchers in a web of unnecessary paper" and that delays in the drug approval process could add millions of dollars to the research and development costs for each successful drug approval by the Federal Drug Administration. The American Association of Medical Colleges says the effects of the regulations could be "devastating" and some institutions will have to cut back their research programmes considerably. Others complain that such restrictive regulation of research will drive researchers to other countries.

The reporting and record-keeping requirements of the new regulations are extensive. For example, before any research which might involve pain in an animal is carried out, the principal investigator must provide a written assurance that alternative procedures were considered and found to be unsuitable, describing what they were, and also providing assurances that the experiment does not unnecessarily duplicate previous experiments.

Under the new regulations, the location of facilities where animals are housed or used in research must be revealed in the institute's annual report. Researchers say that animal activists will have access to the information and security will be jeopardized, thus increasing costs still further. They also complain that APHIS inspectors will be allowed to make copies of records and take photographs. This could lead to the public revelation of trade secrets.

The regulations for the "psychological well-being" of non-human primates have been strongly criticized. Researchers say forcing primates into groups would be traumatic for the animals and dangerous to humans. Field biologists are urging APHIS to exclude wild animals from the regulations, urging that the mandatory twice-yearly inspections will be logistically impossible.

Christine McGourty

Cambridge law passes unanimously

Boston

THE Cambridge (Massachusetts) City Council has ended, at least for the time being, a three-year battle between local research scientists and animal welfare advocates by passing unanimously the nation's first city ordinance to protect laboratory animals. As expected, the vote, taken at the end of June, establishes a "Commissioner of Laboratory Animals", who will have the right to conduct surprise inspections of research facilities within the Cambridge city limits which use animals.

Also as expected, the new law requires all research institutions in the city — including commercial laboratories — to conform to federal statutes and regulations concerning the care of laboratory animals and expands those regulations to include rodents, birds, fish, reptiles and amphibians which are at the moment exempted from the federal Animal Welfare Act. And finally, the law requires all research institutions to create autonomous animal-care committees — with the power to disapprove or restrict experiments — whose membership must include one public member subject to city approval (see *Nature* 339, 496; 15 June 1989).

But the council rejected a key amendment that would have required the presence of an animal-rights advocate on each research institution's animal-care committee. Of all the provisions under consideration, this was the one most vociferously resisted by the research community. David Nathan, chief physician at Children's Hospital and president of the research-orientated group Citizens United for Research and Education, says "the council was wise enough to see that individuals who are morally opposed to any animal research cannot possibly regulate its quality".

Animal-rights supporters saw a victory in the precedent set by the new law and because of the size of the animal population affected. Cambridge, a city with one of the highest concentrations of research laboratories in the United States, is home to thirteen large research institutions which in total use an estimated 60,000 animals a year.

But researchers in the city say that they can live with the city council's decision. Richard Taylor, professor of biology at Harvard University and head of its animal care committee says that researchers are doing "an absolutely first-rate job" and can withstand "any level of scrutiny". According to Taylor, the research community would not tolerate "somebody who's philosophically opposed to research on our committee who could stop the research". After the vote, Taylor spoke for many researchers when he said "it could have been a lot worse".

Seth Shulman

URANIUM

Problems in Hungary

London

HUNGARY'S uranium industry is in trouble. Closure of the country's only uranium mine, in the Mecsek hills, is virtually inevitable, according to industry minister Ferenc Horvath, although the government will not reach a final decision until September.

Under an agreement with the Soviet Union dating from the mid-1950s, the Hungarian government is now allowed to publish details of its uranium production, but it is no secret that the mine is uneconomic. Annual government subsidies to uranium mining have almost doubled since 1985 and this year amount to 2,500 million forints (\$40 million). According to the Hungarian economic journal *Világgazdaság*, if the mine remains open, this subsidy could reach 5,000 million forints by the mid-1990s.

Under the existing agreement, all Hungarian uranium ore has to be delivered to the Soviet Union which supplies fuel elements to Hungary for the 1760-MW nuclear power station at Paks. The intra-Comecon pricing system is based on world market prices averaged over the past five years.

With the current depressed state of the uranium market, the Soviet Union, says *Világgazdaság*, is now paying considerably more than the free market price for yellowcake (\$60–70 as against \$40–45 a kilogram). Furthermore, the Soviet Union allocates to Hungary interest-free loans for geological surveys and also supplies goods — including timber and chemicals — related to the uranium industry.

The Paks power station itself runs at an

apparent profit, contributing an annual 500 million forints to the state budget. But this is calculated on the basis of a generating cost of 1.10 forints per kwh, of which the uranium costs a nominal 0.08 forints. Several months ago, the Paks management was worried about the subsidies to uranium mining being withdrawn and having to buy fuel elements at full production cost. Secrecy restrictions made it difficult for them to draw up contingency plans although, it was hinted, they did not rule out the possibility of buying their fuel elements from some source other than the Soviet Union in the future.

According to Gyula Czipper, the FRENCH RESEARCH

deputy industry minister, at the present rate of production, Hungary's uranium reserves would last at least another 30 years. The deposits, however, which lie at 800–1,000 metres depth, seem likely to remain unprofitable. The management of the Mecsek mine has begun a survey to try to locate small deposits nearer the surface, which would be considerably cheaper to work, but this seems to be a last-ditch stand to stave off closure.

According to *Világgazdaság*, the mine management has now been authorized by the Hungarian government to seek a buyer on the world market, despite the agreement with the Soviet Union. And it now seems likely that the secrecy surrounding the Hungarian uranium industry will be lifted in the next few weeks. **Vera Rich**

Benveniste under review

Paris

DR Jacques Benveniste, whose research on the 'memory of water' put him (and *Nature*, which published his results) at the centre of international controversy last year, has been reprimanded by the French national institute of health and medical research (INSERM). While the reports of two evaluation committees were generally favourable towards the overall work of his laboratory, Unit 200, he has been given a six-month probationary period to re-

said that he would be willing to obey. His research was then re-evaluated by INSERM's scientific council, which included two experts from foreign institutions.

On Monday of this week, Benveniste met INSERM director Phillipe Lazar to hear the final decision about his post as director of research and the future of his research unit. In a statement to the press, the INSERM directorate explained that the two evaluation committees shared a "very favourable opinion of the overall activities of the laboratory", but were "extremely reserved regarding the studies of high dilutions". The statement went on to criticize Benveniste for "an insufficiently critical analysis of the results he reported, the cavalier character of the interpretations he made of them, and the abusive use of his scientific authority *vis-a-vis* his informing of the public".

Benveniste's research unit of immunopharmacology of allergy and inflammation has been given approval to continue to 30 June 1992. Benveniste should also be allowed to stay as director of the unit when his present mandate expires at the end of the year. But this decision is suspended until 31 December "at the latest", during which time he is asked to adopt the "critical and reserved attitude" expected of a scientist.

Speaking to *Nature* before his interview with Lazar, Benveniste remained adamant that he has been unfairly treated and that *Nature* will end up looking foolish for having debunked his research. He pointed out that he has published 67 papers and major reviews in the past four years, of which two are "citation classics". He claimed that his research findings have now been independently confirmed and said that he is "flabbergasted at the risk" taken by *Nature* last year. "What if I am right?", he asked. For INSERM, the risks are the same.

Peter Coles

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

ELECTIONS

Success for Greens

Sydney

THE Greens won their first big victory in Australia last month when they joined a coalition with the Labor party to take power in Tasmania.

The Liberal party won the largest number of seats in the elections in May but with 17 seats were one short of an absolute majority. At the end of June they lost a no-confidence vote, allowing the Labor party, with 13 seats, and the Green Independents, with five seats, to form a coalition government.

The accord between the two parties limits the state's woodchip quota, creates new parks, adds to the list of proposed world heritage sites, prohibits several forest product ventures and transforms the logging industry. Labor is being pushed by the Greens to introduce aboriginal freehold land rights, price controls, equal opportunity laws, four-year parliamentary terms, freedom of information laws and increased funding for education.

Tania Ewing

Beneviste — claims for dilute water

establish "the full guarantee of his peers".

The evaluation of laboratories is a standard four-yearly procedure within INSERM and was not provoked by the controversy surrounding Benveniste's research. But when the first review by a specialist pharmacology committee was carried out in April, there were already signs that his work on high dilutions would not be judged favourably. They "issued a statement asking me to stop work on high dilutions, without any scientific reasons", he told *Nature*. Benveniste protested, but

Approval hoped for soon

Washington

A CONTROVERSIAL survey on adult sexual habits in the United States, aimed principally at helping epidemiologists to predict how the AIDS virus will spread, is moving into the open again, having been stifled for months by opposition from within the Bush Administration. The survey, which includes questions on homosexual and bisexual practices, has already passed through a long bureaucratic chain from its designers at the National Institute of Child Health and Human Development (NICHD) to the White House Office of Management and Budget (OMB) and back again. If the wrath of conservative congressmen such as William Dannemeyer (Republican, California) can be avoided, an expurgated version of the survey may soon come up for final approval.

Opposition to the survey has been channelled through OMB, which oversees all federally supported surveys. OMB director Richard Darman wanted a review of the survey to ensure that every question had a demonstrable purpose for AIDS researchers and that no unnecessary private questions were included.

He passed his request to Louis Sullivan, secretary of the Department of Health and Human Services. Sullivan passed the task to James Mason, his assistant secretary of health, who passed it to James Wyngaarden, director of the National Institutes of Health (NIH), who passed it to NICHD, which

conducted an internal review as well as asking for the advice of its independent advisory panel. NICHD decided, without involving the scientists who had put together the original survey, to delete some of the questions. Out went inquiries about masturbation and sexual fantasies, which were deemed to have little relevance to the spread of AIDS. The panel unanimously endorsed the changes at its June meeting, since when the survey has worked its way back up the paper chain as far as Mason's desk, Wyngaarden having given it the *imprimatur* of NIH.

The survey is a pilot study of 2,300 people, costing about \$2 million, which is designed to help prepare a later full-scale survey. The National Opinion Research Center (NORC), a non-profit social sciences organization affiliated to the University of Chicago, has already been awarded the contract to conduct the pilot survey, which will consist of personal interviews, both in homes and in a research environment, and telephone interview.

Potential interviewees will be selected from lists of people who have already participated in public-health surveys on other matters. The selection will not be completely random, because the survey designers want to ensure that at-risk minorities, selected by age, marital status and ethnicity, for example, are adequately represented. Trained interviewers should be able to administer the survey in about an hour.

Among the questions that have caused problems are those on the prevalence of past and present homosexuality in married men, and heterosexual anal intercourse. Researchers say they are essential because the poor quality of available data, much of which dates back to the Kinsey report of the 1940s, prevents clear predictions of the future incidence of AIDS in the population at large. But opponents regard such questioning as intrusive when asked of people who are not obviously in any high-risk part of the population.

Dannemeyer, a dedicated conservative, has been the most vocal congressional critic of the survey. He believes such surveys are fundamentally unable to answer the questions they pose because people's willingness to respond depends on their proclivities. Dannemeyer favours instead a programme of confidential reporting of AIDS cases and tracking of sexual partners. He has no objections to asking specific questions about recent sexual contacts of people with AIDS; he and his colleagues oppose not so much the questions themselves as their being asked in what they regard as an indiscriminate manner.

Few congressmen have come out vocally in favour of the survey, although Representatives Bill Green (Republican, New York) and Henry Waxman (Democrat, California) have signed a letter of commendation to Sullivan. Supporters of the survey believe most politicians are on their side, but are nervous of saying so in an era when President Bush was elected on a wave of support for "family values".

NIH officials feel that once Sullivan gives his approval, the survey can go ahead. In a letter to Sullivan, Darman implied that his office has no legal power to prevent the survey, provided it passes proper scrutiny, and NIH officials are hoping that the pilot survey can go ahead without any public commotion.

But Dannemeyer is watching and, according to a spokesman, is ready to tell Darman the survey would be an embarrassment to the Bush administration. Failing that, he will try to have the process halted legislatively in the Senate, as the House of Representatives does not have that power. As the close-mouthed supporters of the survey are well aware, public attention is more important than legalistic nicety when such a charged issue is caught in the spotlight. **David Lindley**

■ A plan to survey the sexual practices and attitudes of 20,000 UK citizens to help predictions of the spread of the AIDS virus is under discussion following the successful completion of pilot studies. The survey is designed by Social Community Planning Research, the organization that would conduct it, together with researchers at Imperial College and the medical schools of St Mary's and Middlesex Hospitals, all in London. **P.N.**

Old morality study finally surfaces

Washington

WHILE scientists struggle to set in motion an up-to-date survey of the sexual habits of people in the United States (see above), they can at least celebrate the belated publication of the results of a survey carried out almost 20 years ago. The data, which languished unseen because of a squabble over the manner of publication, reveal a population possessed of rather stern opinions: 60 per cent think that any kind of homosexual behaviour should be a crime, and 25 per cent would like to see laws prohibiting adultery.

"Sex and Morality in the US", published by the Wesleyan University Press, describes the beliefs and habits of a sample of 3,018 people who were questioned in late 1970. The study was performed for the Kinsey Institute for Research in Sex, Gender, and Reproduction in an attempt to extend the Kinsey reports of the 1940s and 1950s. Soon after the data were collected, some abbreviated accounts of the results emerged, but the details of the survey remained mired in a lengthy analysis which turned into a quarrel over the arrangement of authors' names and royalties. The study was published only in response to pressure

from people who saw its potential value for AIDS research.

The study reveals something of a gap between what people think they ought to be doing and what they actually do. More than half the respondents declared that premarital sex is wrong, but almost 80 per cent had had intercourse with someone other than their spouse before they married, and 60 per cent expressed no regret about having done so. And although 60 per cent of Americans apparently think homosexuality should be outlawed, more than 50 per cent say they would remain friends with people they discovered to be homosexual.

The immediate relevance of the data to estimates of the spread of AIDS is questionable, because everyone in the survey was at least 21 years old in 1970, long before the first cases of AIDS were recognized. The survey makes no attempt to determine the prevalence of practices that might be involved in the transmission of AIDS. For example, 17 per cent of men and 9 per cent of women said they had experienced some kind of homosexual activity, but it is not apparent what proportion of these respondents were simply confessing to teenage experimentation. **David Lindley**

The benefits of the commons

F. Berkes, D. Feeny, B.J. McCay and J.M. Acheson

Conventional wisdom holds that resources held in common will invariably be overexploited — the “tragedy of the commons”. A number of examples show that this is not necessarily so.

It has become a truism that resources held in common are vulnerable to over-exploitation. Twenty-one years ago, Hardin popularized this dilemma — calling it the “tragedy of the commons” — by the use of a metaphorical village common in which each herdsman “is locked into a system that compels him to increase his herd without limit”¹. Hardin argued that such problems have no technical solutions, and emphasized the need for government controls to limit “freedom in the commons [which] brings ruin to all”¹. Hardin and others² have subsequently pointed to privatization of common resources as another solution consistent with the analysis of many resource economists³.

It is usual to assume that resource degradation is inevitable unless common property is converted into private property or government regulations are instituted. The prevalence of this view is reflected by an article in *The Economist* of 10 December 1988 about fisheries, typically viewed as a common-property resource: “...it is possible to manage fisheries successfully”, the author asserts, “provided three facts are kept in mind”. Two of these are relevant here: “left to their own devices, fishermen will over-exploit stocks” and “to avoid disaster, managers must have effective hegemony over them”.

Nevertheless, research carried out in the 21 years since Hardin’s article often leads to conclusions that challenge this conventional wisdom. Such results are of interest to resource managers, applied natural and social scientists, policy-makers and development planners. Many case studies, including our own, show that success can be achieved in ways other than privatization or government control⁴⁻⁷. Communities dependent on common-property resources have adopted various institutional arrangements to manage those resources, with varying degrees of success in achieving sustainable use. We use ecological sustainability⁸ as a rough index of management success without necessarily implying resource use that is ecologically or economically optimal.

As a first step in the analysis, it is necessary to define the kind of resources under consideration. Common-property (or common-pool⁹) resources share two key characteristics. First, these are resources for which exclusion (or control of access) of potential users is problematic. The physical nature of the resource is such

that controlling the access of potential users is costly and, in some cases, virtually impossible. Migratory or fugitive resources such as fish and wildlife pose obvious difficulties. Similarly, ground water, range and forest lands, and global commons⁸ such as the high seas, the atmosphere and the geosynchronous orbit, pose problems of exclusion.

The second key characteristic of

right to exclude others from using the resource and to regulate its use. (3) Under communal property, the resource is held by an identifiable community of users who can exclude others and regulate use. Some shellfish beds, range lands, forests, irrigation and ground water have been managed as communal property. (4) State property or state governance means that rights to the resource are vested



Cree Amerindian fishermen of James Bay, seining river eddies for whitefish. The use of the resource is regulated under rules agreed upon by all — groups of fishermen wait their turn for the best sites during the short fishing season. (F. Berkes.)

common-property resources is subtractability; each user is capable of subtracting from the welfare of others. This characteristic creates a potential divergence between individual and collective economic rationality in joint use³. As one user continues to pump water from an aquifer, others experience increased pumping costs; as the number of fishing boats increases, the catch per unit of effort for each declines. On the basis of these two characteristics, we define common-property resources as a class of resources for which exclusion is difficult and joint use involves subtractability.

As a second step in the analysis, a taxonomy of property-rights regimes is needed⁹⁻¹¹. Common-property resources are held in one of four basic property-rights regimes. (1) Open access is the absence of well-defined property rights. Access is free and open to all, as with ocean fisheries of the past century. This is the regime implied in Hardin’s model. (2) Private property refers to the situation in which an individual or corporation has the

exclusively in government, which controls access and level of exploitation. Examples include crown lands and resources such as fish and wildlife held in public trust. These four categories are ideal, analytical types. In practice, resources are often held in overlapping combinations of these four regimes, and there is variation within each.

We now briefly summarize selected case studies. These studies show the workings of communal-property systems not recognized in Hardin’s model, as well as the limitations to the use of state governance in some situations.

Our first case concerns wildlife hunting territories in James Bay, Quebec, in northeastern Canada¹². Hunters in this subarctic area have traditionally used resources communally, as do many Amerindian groups, and have a rich heritage of customary laws to regulate hunting. Beaver is an important species both for food and, since the start of the fur trade in James Bay in 1670, for commerce. ▶

The beaver is vulnerable to depletion because colonies are easily spotted. A community-based hunting territory system, with senior hunters and their families acting as stewards of specific territories, at present ensures sustainable use. The beaver resource in James Bay, however, has not always been used sustainably. In the 1920s, a large influx of non-native trappers followed the new railroad into the area to take advantage of high fur prices. Amerindian communities lost control over their territories and all trappers, including natives, contributed to a "tragedy of the commons". Conservation laws were eventually enacted after 1930, when beaver populations were at an all-time low, and outsiders were banned from trapping in James Bay. Amerindian community and family territories were legally recognized and customary laws became enforceable, resulting in productive harvests after about 1950¹². The experience of the 1920s and 1930s is not unique. Periods of cut-throat rivalry among fur companies had led to non-sustainable use of resources twice before: in the mid-1700s and in 1825–29. Gradually, however, local control was restored and stocks recovered¹².

Our second and third cases deal with lobster and fish management on the east coast of the United States^{13,14} and show that communal territories exist even in societies that subscribe to the ideal of freedom in the commons. In the US tradition, marine resources belong to all citizens but are controlled by state governments as a public trust. Privatization of some marine resources such as shellfish beds is feasible but not always socially desirable or politically acceptable¹⁵. Government management is similarly difficult: limiting the number of licences is considered an infringement of citizens' rights. Even so some groups of users are able to restrict access and manage common-property resources.

The lobster resource is vulnerable to overharvesting, but lobster stocks in Maine have remained sustainable. Although some managers have for decades been predicting a resource collapse, the Maine lobster catch has been remarkably stable since 1947¹³. The state government establishes lobstering regulations but does not limit the number of licences. In practice, however, there is exclusion through a system of traditional fishing rights; to go lobster fishing at all, one has to be accepted by the community. Once accepted, a lobsterman is only allowed to fish in the territory held by that community. Interlopers are usually discouraged by surreptitious violence.

One cannot say if the resource could

have been used sustainably in the absence of such locally enforced exclusion and regulation. But we have compared the productivity of exclusively used territories with areas in which claims of adjacent communities overlap. We found that fishermen in the exclusive territories catch significantly more and larger lobsters with less overall effort¹³.

The third case, a trawl fishery in the New York Bight region, provides an alternative community-based solution to the commons dilemma¹⁴. The fishermen who belong to a cooperative specialize in the harvest of whiting. They have ready

Beaver hunters in north-eastern Canada (this one was sketched in 1891) used resources communally until the coming of the railroad in the 1920s brought an influx of 'outsiders'.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Mary Evans

was essentially unregulated. The rapid commercial exploitation of teak in Thailand in the late nineteenth century led to the nationalization of all forests. State ownership fails to provide consistent enforcement, but it also serves to deny users the authority to manage local forests. Illegal logging, followed by further land clearing for cultivation, is widespread. Although much of this land is suitable for cultivation, there are few safeguards for conserving environmentally sensitive areas; this results in overall damage to land.

The lack of enforcement of state-forest property rights leading to accelerated degradation is not unique to Thailand. The nationalization of forests in Nepal (1957) and Niger (1935) produced a similar outcome¹⁷. In Nepal, the situation is being ameliorated by the re-creation of communal management at the local level¹⁸. Without effective control by government, nationalization has often converted traditional communal property into *de jure* state property but *de facto* open-access.

Having reviewed a few cases, we return to the tragedy of the commons model to explore its problems in relation to the findings. Hardin asks the reader to assume a pasture "open to all"¹. Each herdsman acts in an individually rational fashion by adding animals to the common pasture. For him, the private benefits of adding one more animal exceed the private cost. Because each herdsman does the same, the overall result is overgrazing and disastrous losses for all.

Hardin's model provides insight about the divergence between individual and collective rationality. But it fails to take into account the self-regulating capabilities of users. It assumes that the herdsmen are unable to limit access or institute rules to regulate use. Therefore, overexploitation is inevitable — unless privatization or government controls are imposed. These conclusions have been used as part of the justification for nationalization¹⁸, privatization of land resources¹⁹, and the widespread practice of top-down development planning that ignores local institutions^{4,6}. The social and ecological costs of these practices have often been tragic in their own right.

Recognition that users have the potential and, under some conditions, the motives and means to act collectively opens up other policy alternatives and provides questions about why some communal management systems fail and others succeed. The success or failure of common-property resource management has to do with the exclusion and regulation

access to the best whiting grounds in the region, and often dominate the regional whiting market in the winter months.

The cooperative maintains relatively high prices for members through supply management; it limits entry into the local fishery and establishes catch quotas among members. Limited entry is achieved through a closed membership policy and the control of docking space, effectively excluding non-members from access to whiting grounds and markets. Quotas are based on the estimates of what the cooperative can sell to the regional market, and are achieved in ways that reward individual initiative but also discourage 'free-riding'. By contrast with government-imposed regulations, which are considered by fishermen to be inflexible and which in any case are ineffective because they do not address the fundamental problem of access, self-regulation through the cooperative is considered to be both flexible and effective in maintaining sustainable use¹⁴.

Forests in Thailand comprise our fourth case¹⁶. Traditionally the exploitation of high-value timber was regulated by local governments; the use of low-value timber

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Woodcutters near Bhartang, Nepal — nationalization of Nepal's forests led to over-exploitation, but the situation is now being improved by the re-creation of local communal management.

of joint use. Forest destruction in Thailand, for example, occurs because villagers do not own the forest and cannot exclude others. Local people therefore have little incentive to conserve and every incentive to cut down trees before someone else does¹⁶.

By contrast, in other examples — hunters in James Bay, lobstermen in Maine, trawlermen in the New York Bight area, communal forest users in Nepal, and irrigation water users in South India²⁰ — groups are able to exclude other potential users and regulate their own joint use. They are therefore able to reap the benefits of their own restraint. Our examples are not isolated, but are consistent with a large body of literature on grazing lands²¹, forests²², water²³ and coastal marine resources²⁴, covering a wide range of regions and cultures throughout the world.

What accounts for the many exceptions to the predictions of the conventional theory? How can Hardin's model be improved to obtain a more comprehensive theory of common-property resource management? First, the Hardin model confuses common-property resources with open access — the absence of property rights. By equating common-property resources with open access, and then assuming that open access leads to overexploitation, the model falls into the trap of equating the commons with over-exploitation.

Second, the model assumes that the individual interest is unconstrained by existing institutional arrangements. In many communities, common-property resource users are compelled by social pressure to conform to carefully prescribed and enforced rules of conduct.

Third, the model assumes that resource users cannot cooperate toward their common interests. This is not necessarily so; under certain circumstances, voluntary

collective action is feasible²⁵, and sustainable outcomes are not unusual^{4-7,20-24}.

More fundamentally, the model overlooks the role of institutions that provide for exclusion and regulation of use. Cultural and historical factors underlying such institutional arrangements are a key to the success of communal management of coastal marine resources in Japan and several Pacific-island nations²⁴, in addition to the cases we describe above.

Finally, the set of solutions offered by the model is too limited. Privatization or the imposition of government control are not the only viable policy options. In fact, the conventional reliance on these approaches is overly sanguine. By definition, common-property resources are ones for which exclusion is difficult and so privatization is often not feasible. Although dividing a commons and assigning individual property rights can increase efficiency under some circumstances, it might not in others. Similarly, state control

has worked in some cases, but the example of Thailand forests illustrates its potential for failure.

In general, we propose that successful approaches to the commons dilemma are found in complementary and compatible relationships between the resource, the technology for its exploitation, the property-rights regime and the larger set of institutional arrangements. We also propose that combinations of property-rights regimes may in many cases work better than any single regime. The success of local-level management, for example, often depends on its legitimization by central government; James Bay¹² and recent experience in Nepal¹⁸ are examples. Such nested relationships are also found in fisheries in Japan and Oceania²⁴. In some cases, cooperative management arrangements (co-management) are needed, involving the sharing of power between governments and local communities²⁶.

In sum, sustainable common-property resource management is not intrinsically associated with any particular property-rights regime. Successes and failures are found in private, state and communal-property systems. Recent research highlights the potential viability and continued relevance of communal-property regimes, nested systems and co-management. Studies after that of Hardin have shown the dangers of trying to explain resource use in complex socio-ecological systems with simple deterministic models. □

F. Berkes is in the Institute of Urban and Environmental Studies, Brock University, St Catharines, Ontario, Canada L2S 3A1; D. Feeny is in the Department of Economics, McMaster University, Hamilton, Ontario, Canada L8S 4M4; B. J. McCay is in the Department of Human Ecology, Cook College, Rutgers University, PO Box 231, New Brunswick, New Jersey 08903, USA; J. M. Acheson is in the Department of Anthropology, University of Maine, Orono, Maine 04469-0158, USA.

- Hardin, G. *Science* **162**, 1243–1248 (1968).
- Hardin, G. & Baden, J. (eds) *Managing the Commons* (Freeman, San Francisco, 1977).
- Gordon, J. S. *J. polit. Econ.* **62**, 124–142 (1954).
- Proc. NRC Conf. *Common Property Resource Management* (National Academy, Washington, DC, 1986).
- McCay, B. J. & Acheson, J. M. (eds) *The Question of the Commons: The Culture and Ecology of Communal Resources* (University of Arizona Press, Tucson, 1987).
- Ostrom, V., Feeny, D. & Picht, H. (eds) *Rethinking Institutional Analysis and Development* (Institute for Contemporary Studies, San Francisco, 1988).
- Berkes, F. (ed.) *Common Property Resources: Ecology and Community-Based Sustainable Development* (Belhaven, London, 1989).
- World Commission on Environment and Development *Our Common Future* (Oxford University Press, 1987).
- Ostrom, E. in Proc. NRC Conf. *Common Property Resource Management* 599–615 (National Academy, Washington, DC, 1986).
- Bromley D. W. in Proc. NRC Conf. *Common Property Resource Management* 593–598 (National Academy, Washington, DC, 1986).
- Bromley, D. W. *Economic Interests and Institutions* (Blackwell, Oxford, 1989).
- Berkes, F. in *Common Property Resources: Ecology and Community-Based Sustainable Development* 70–78 (Belhaven, London, 1989).
- Acheson, J. M. *The Lobster Gangs of Maine* (University Press of New England, 1988).
- McCay, B. J. *Anthrop. Q.* **53**, 29–38 (1980).
- McCay, B. J. in *The Question of the Commons: The Culture and Ecology of Communal Resources* 195–216 (University of Arizona Press, 1987).
- Feeny, D. in *World Deforestation in the Twentieth Century* (Duke University Press, 1988).
- Thomson, J. T., Feeny, D. H. & Oakerson, R. J. in Proc. NRC Conf. *Common Property Resource Management* 391–424 (National Academy, Washington, DC, 1986).
- Arnold, J. E. M. & Campbell, J. G. in Proc. NRC Conf. *Common Property Resource Management* 425–454 (National Academy, Washington DC, 1986).
- Peters, P. in *The Question of the Commons: The Culture and Ecology of Communal Resources* 171–194 (University of Arizona Press, 1987).
- Wade, R. *Village Republics: Economic Conditions for Collective Action* (Cambridge University Press, 1987).
- Netting, R. M. *Balancing on an Alp: Ecological Change and Continuity in a Swiss Mountain Community* (Cambridge University Press, 1981).
- Fortmann, L. & Bruce, J. W. (eds) *Whose Trees? Proprietary Dimensions of Forestry* (Westview, Boulder, 1988).
- Maass, A. & Anderson, R. L. *... and the Desert Shall Rejoice* (MIT, Cambridge, 1978).
- Ruddle, K. & Akimichi, T. (eds) *Maritime Institutions in the Western Pacific* (National Museum of Ethnology, Osaka, 1984).
- Ostrom, E. in *Rethinking Institutional Analysis and Development* 101–139 (Institute for Contemporary Studies, San Francisco, 1988).
- Pinkerton, E. (ed) *Co-operative Management of Local Fisheries* (University of British Columbia Press, 1989).

India's too costly imports

SIR—I write in support of the views of Patrick J. Willems in his note on 'Primer availability' for his prenatal diagnosis work (*Nature* 337, 10; 1989) and I would like to add to them. I am sure that my counterparts in this country and in other countries in this part of the world face similar problems in pursuing their research in molecular biology.

Some United States-based companies charge us very high prices for molecular biology reagents—far more than the true catalogue prices. I calculate that if I were working in a US laboratory, I would receive enzymes and reagents at half the price I have to pay in India. Whenever we ask for quotation, the prices quoted are one and a half times greater than catalogue prices. But the extra handling charges are exorbitant, sometimes half of the total cost. We also pay insurance charges, customs charges, freight charges and local agents' commission.

Laboratory instrumentation is another problem. One US company quoted \$120,000 to us for an automatic nucleotide sequencer sold in the United States for \$90,000. When I sought an explanation from one of the company's managers (who was visiting India to demonstrate the instrument) the answer was highly unsatisfactory.

These companies even give a discount in the United States on their catalogue prices. Here, we are willing to pay all their extra charges yet we are forced to pay over and above the catalogue prices. What for? We already suffer from procedural delays when trying to obtain equipment from abroad, so an item which one can procure in a US laboratory on the same day, or the next morning, takes several months to reach an Indian customer. Yet we maintain our good work, which is quite competitive, even though, sadly, some foreign journals show bias in accepting our results. So a scientist in this part of the world suffers from both sides.

Will the companies not change their attitude and look into the matter more seriously for the betterment of science, which is serving humanity all over the world?

Y. D. SARMA

Department of Biotechnology,
All India Institute of Medical Sciences,
New Delhi – 110 029, India

Food comes first

SIR—In return for the 'India in France' festival, France is constructing in New Delhi an Indo-French Centre for Advanced Research as part of the 'France in India' festival. It is hoped that this institution will be an effective centre for excellence in

scientific research and will not fall prey to the maladies gripping other such establishments in India, where science and research are in jeopardy.

India has a network of research laboratories, some well-equipped even by European standards. But for reasons debated earlier in these columns, most research centres in India, in spite of the availability of capable men and sufficient materials, have lagged behind in contributing their share to scientific growth.

Indian scientific circles are aware of this fact but no effort has been made to root out retrogressive forces impeding adequate scientific progress in India, where sycophancy rules supreme. Several commissions have been instituted to recommend remedial measures, but the fate of these recommendations is usually to end up gathering dust on government shelves.

The Centre for Advanced Research which France is building must be made to run as a first-class research centre under French guidance before being handed over. Otherwise there is a danger that the proposed centre will join dozens of other research centres where excellence in science is sacrificed to vested interests.

In France there exists an efficient means of producing milk and cheese at village or departmental level. If this system could be demonstrated in India during the 'French Festivities' as a model, it would be of practical value for the nutrition of the Indian masses, 80 per cent of whom live on a vegetarian diet. Agricultural production and nutrition deserve priority in Indo-French ventures.

A. N. MALVIYA

CNRS,
5, rue Blaise Pascal,
67084 Strasbourg Cedex,
France

Cry from the past

SIR—While sorting through the papers of the late Martin D. Burkenroad held in the archives of the San Diego Society of Natural History, the following letter was found.

January 10th, 1939

Prof. Williard G. Van Name
American Museum of Natural History
New York City, USA

Dear Sir,

I hope that you will forgive me for taking the liberty of writing to you in this way. I have been working as an assistant at the Istituto di Anatomia Comparata della R. Università di Napoli for five years.

Last month, owing to the laws against the Jews, I lost my position and since there are no further prospects here for me, I would be really much obliged to you if you would let me know whether there are any possibilities of obtaining employment in the United States of America.

I am a Doctor of Natural Science, twenty five years of age, with a fair knowledge of English,

French, Russian and, of course, Italian. I have had a thorough training in Zoology and Comparative Anatomy and I have published twelve scientific papers up to the present. I have devoted myself to biological research, and have also had considerable experience in training and organizing.

If there is any possibility in the United States of getting a technical position in Biology, I am quite ready to work as a secretary: in fact, I have also had much experience in bookkeeping, typing, and general office routine.

I most sincerely hope that you will excuse me for all the trouble to which I am putting you.

Thanking you in anticipation.

I am

Yours very faithfully

(Signed) Isabella Coifmann

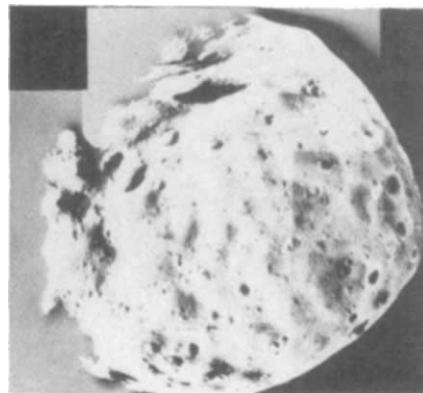
Apparently, Prof. W.G. Van Name had forwarded the letter to Burkenroad in the hope that there might have been an open position at Yale's Bingham Oceanographic Laboratory. That the department volunteer who was assisting me with the sorting of Burkenroad's papers, and who brought this item to my attention, is himself Jewish, added to the letter's poignancy. Besides sharing this footnote in history with the readers of *Nature*, we at the San Diego Natural History Museum would like to know whatever happened to Dr Coifmann. If someone has any information in this matter, please contact us.

FREDERICK R. SCHRAM

Department of Paleontology,
San Diego Natural History Museum,
PO Box 1390,
San Diego, California 92112, USA

Mistaken identity?

SIR—The photograph issued by NASA purporting to be a photomosaic of Phobos (*Nature* 339, 9; 1989) bears a quite remark-



Assuredly a martian moon.

able resemblance to a fried spicy poppadam (samples of which are available in all good Indian take-aways). Is fraud to be suspected?

LAURENCE D. HURST

Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK

Do authors need guidance?

The wish that news of discovery should quickly reach the public that pays for it should be more often tempered by deliberation. Prior publication is, by definition, premature.

IT is well-known that there is nothing so boring as when journals like this write about themselves or about others that resemble them, but it may not yet be time to switch to another page. What follows is about the guidance given by journals other than this to researchers who are would-be authors. That guidance, to say the best of it, is confusing, but no more so than in the matter of the propriety of prior disclosure. The occasion for what may seem too much like a homily is the appearance in *Physical Review Letters*, and also in other journals from the US Institute of Physics, of what may seem to be an all-embracing, or even over-embracing, catalogue of dos, don'ts and maybes.

PRL, as it is known, is among the most authoritarian of all journals. It is especially dogmatic on the space that contributions are allowed to fill. At regular intervals, there is published a standard recipe by which contributors can calculate for themselves how many keystrokes by their typist's fingers will fill up the maximum of four pages of text. Headings, of course, count extra, references must be tallied separately and there is a kind of algorithm for working out the space occupied by diagrams. It is remarkable how well contributors respond: almost all published articles turn the third page, but the upper limit is evidently enforced.

PRL is also among the most authoritative of journals, which attribute derives from its creator and first editor, the late Sam Goudschmidt (with Uhlenbeck, the premature discoverer in 1926 of electron-spin). Goudschmidt's cleverness was to recognise that a few dedicated typists working with typewriters equipped with interchangeable fonts on golf-balls could come near to simulating typesetters' type, but could also send a contributor a simulation of his typeset proof more quickly than conventional journals. (The need to simulate has now, of course, disappeared.) Goudschmidt's personal concern for the integrity of what his journal published, matched only by Chandrasekar's care for the *Astrophysical Journal*, was naturally another important consideration.

As it happens, Goudschmidt during his life had only one consistent complaint against his authors — their evident liking for being published on the front page of the *New York Times* as well as in *PRL*. Yet in the 1960s, the habit seemed incorrigible: there was always an anti-proton or some other exotic particle whose dis-

covery seemed that important. Goudschmidt's war against the practice of prior release was so effective that, even if science never disappeared for long from the front page of the *New York Times*, at least his contributors became well-versed in explanations of how their work had arrived there by accident.

How things have now changed. *PRL*'s latest version of its authors' guide, bound in with its issue of 3 July, includes a statement on prior disclosure that will have even unshockable Sam Goudschmidt turning in his grave. The journals of the American Physical Society, the statement says, are just like other journals in requiring that authors should not previously have published what they have to say elsewhere. The explanation, the statement says, is that, "for decades", editors "have frowned on multiple publication". (Presumably, if they chose, they could stop frowning, and maybe some have stopped already.)

But there are naturally exceptions to every rule — abstracts prepared for meetings and later published, students' theses and even "abbreviated or preliminary" accounts of research appearing in conference proceedings. Authors are warned that it is their job to get permission for reproducing duplicated material from whoever may hold the copyright.

Curiously, *PRL* does not on this occasion refer to the most abundant form of duplicate publication — the preprint business. Yet *PRL* functions in many of the fields in which preprints are most common. In high-energy physics, for example, active researchers have been exchanging advance copies of what they plan to send to *PRL* since long before the journal itself came into being, largely because there were then no journals rapid enough to accommodate their needs. But now that *PRL* has been established as both rapid and catholic, instead of the preprint habit dying away, it has spread to other important fields of physics.

The explanation is straightforward. While a preprint may outwardly seem a means by which innocent researchers tell their competitors what they are about sooner than would otherwise be possible, they are also effective ways of ensuring that everybody knows who did what first. Informally, preprints and not publications determine the priority for discovery. But, sadly, preprints are innocent of referees' opinions and their consequences, tend to

be laden with unwarranted speculations and are generally confusing. Nobody would now seriously suggest that they should be abolished, for the habit is too deeply engrained, but *PRL* does not even raise the need that they, like students' theses, should be exempted from the interdiction of "multiple publication".

There is worse to come. *PRL*, on behalf of the American Institute of Physics, repeats its now familiar exceptions from the restraints of multiple publication the case in which news of scientific discoveries appear first in popular magazines — *Time*, *Newsweek*, *Scientific American* and *Physics Today* get special mention, but prior publication by "newspaper, television and radio" is not to be held a restraint of publication in one of the journals of the American Physical Society. The reason given is that publication of research should be "timely", especially because "much research is funded by public agencies".

The argument carries little weight. What there is consists of the expectation that fresh news of discovery is more immediate, and therefore more compelling, than news which has been filtered through the peer-review systems of the journals. And on the principle that there must be exceptions to any rule, only a true pedant could insist that news of the discovery of a moon for Neptune should be this week withheld from anybody once it is to hand or, released, should prejudice eventual publication. But there are serious drawbacks in the American Physical Society's encouragement of early popular publication of general news of discovery in science. What, for example, if the authors are afterwards found to be mistaken?

What this implies is that the American Physical Society should think hard about its policy of unrestrained prior release — and then preferably abandon it. Even in cut-and-dried physics, there have been too many unsustainable claims for most people's comfort. But the more serious misfortune of the present arrangements is that the chance that particular discoveries will command their fair share of the general public's attention depends too much on the doings of institution's public relations enterprises — and even those in universities are not above reproach. Sam Goudschmidt, being the first, may have been an unsystematic editor of *PRL*, but his instincts on this issue were correct.

John Maddox

Oldest Eurasian stone tools

Eric Delson

WHEN and where did the first tool-making humans reside on the European continent? This question was addressed at a recent symposium* in Paris and answered in an article published subsequently by Eugène Bonifay¹. Bonifay claims that deposits at St Eble, near the extinct Coupet volcano in the Auvergne region of the French Massif Central, yield quartz fragments which testify to the presence of Palaeolithic humans before 2 million years (Myr) ago.

Flaked stone artefacts are the first clear documentation of human cultural behaviour, and thus the earliest stone tools have always been regarded as a watershed in prehistory. In recent decades, the location of these earliest tools has been assumed to be Africa, particularly eastern Africa, given the completeness of

tion by Dennell *et al.*³ of several purported stone tools from deposits of that antiquity in Pakistan. In this case there have been questions about the age of the horizon, the primary association of the objects with that horizon and the artefactual nature of the objects, not to mention the lack of human or animal fossils around the objects.

The case for St Eble seems to be stronger. During the meeting, participants from various European countries (from Spain to the Soviet Union) and French regions presented information about the earliest archaeological remains in their territory (see ref. 4 for details). Unfortunately, actual examples of these early artefact assemblages could not be provided, and evaluation of the human origin of such objects, especially when made of quartz or

just been published⁵.

But for many the most exciting new locality is St Eble. According to Bonifay¹, the local stratigraphy comprises a gneiss basement overlain by 3–4 m of sandy-clay slope deposit and then about 2 m of bedded volcanic explosion breccias, above which is formed the Coupet volcano itself. On the flank of the volcano is located a younger layer of yellow-orange silt, formed by alteration of the scoria cone and containing a fauna of fossil mammals dating to the middle or late Villafranchian. This local fauna has never been published in detail, nor has the radiometric date of about 2 Myr ago often cited for it — but the latter is in progress by geologists at the University of Clermont-Ferrand. Several artefacts have now been reported from the slope deposit, which could be seen during the excursion to be relatively fine-grained and probably of low-energy origin; Bonifay indicates that no volcanic materials are included within it. If this geological interpretation is correct, the slope deposit is older than 2 Myr and unlikely to yield naturally flaked quartz.

It was not possible to examine the supposed artefacts during the excursion, but I later visited Bonifay's laboratory in Marseilles, where I was shown all available material from this and other localities reported in ref. 5. Several hundred quartz flakes and chunks have been recovered at St Eble, some derived from the basement gneiss and others from heavily rolled pebbles. Bonifay¹ describes five of these as of unquestioned human manufacture and figured two in line drawings. Three are illustrated in the figure, plus one fragment from the fossiliferous silt. These pieces are not obviously human artefacts, at least to one not expert in dealing with quartz implements, but among specialists there appears to be divided opinion. Further details on the fauna and radioisotope age determination, as well as additional archaeological specimens, will be required before a final decision can be reached on the nature of the St Eble assemblage. But for the time being, it is one of the most intriguing of the purportedly ancient vestiges of human occupation in Europe. □

Eric Delson is Professor of Anthropology at Lehman College and the Graduate School, City University of New York and works in the Department of Vertebrate Palaeontology, American Museum of Natural History, New York, New York 10024, USA.



Three purported quartz artefacts from the slope deposits at St Eble in the Auvergne, and one (in the upper centre) from the fossiliferous silts, dating to greater than 2 Myr age. (Photograph by E. D.)

its fossil record². The acceptance of tools in those sites dated to more than 2 Myr ago is based on the objects being clearly identified as human artefacts, their unquestioned dating by chronometric methods and (in most cases) their association either with human fossils or with animal bones which are apparently linked to human activity.

The suggestion of a similar antiquity for stone tools elsewhere in the world, given the lack of human fossils outside Africa older than about 1.2 Myr, immediately elicits reactions ranging from question to disbelief. Such has been the case during the past few years, following the publica-

other coarse-grained materials, is difficult when only drawings are available. Many palaeoanthropologists, including myself, have long questioned any reports of artefacts in Europe older than about 0.5 Myr, at which time the earliest human fossils appear there, but several sites discussed at the meeting appear to document human occupation back towards 1 Myr ago at least.

Another group of even less widely accepted sites ranges back to 2.5 Myr ago. Many of these are located in the Massif Central region of France, and an excursion to study them was organized by Bonifay at the end of the meeting. Most of these localities (for example, Soleihac, Ceyssaguet and Chilhac) have been previously reported, and a brief survey has

* The Earliest Human Occupation of Europe, part of the 114th French National Congress of Learned Societies, Paris, 3–9 April 1989.

1. Bonifay, E. *Comp. Rend. acad. Sci. Paris Ser. II*, **308**, 1567–1570 (1989).
2. Toth, N. & Schick, K. in *Advances in Archaeological Method and Theory* (ed. Schiffer, M.) 1–96 (Academic, London, 1986).
3. Dennell, R. *Curr. Anthropol.* **30**, 318–322 (1989).
4. Ackerman, S. *Science* (in the press).
5. Bonifay, E., Consigny, A. & Liabeuf, L. *Comp. Rend. acad. Sci. Paris Ser. II*, **308**, 1491–1496 (1989).

Mitogenic neurotransmitters

Michael R. Hanley

DESPITE significant progress in the description of vertebrate neural cell lineages and differentiation, little is known about how the extensive proliferation of early brain-precursor cells is coordinated, or how non-neural cells are stimulated to divide during development or following injury. The question of neural cell mitogens has focused on the actions of known growth factors, but the report of Capon and colleagues on page 146 of this issue¹ calls attention to a possibility that is gaining support; that neurotransmitters themselves have the potential to act as regulators of nerve cell growth and division.

Using both transfected Chinese hamster ovary cells and various cell lines expressing identified muscarinic acetylcholine receptor subtypes, the authors note that there is a strict correspondence between the occurrence of specific muscarinic receptors and the relative ability of the stable acetylcholine analogue carbachol to drive DNA synthesis. Specifically, two muscarinic receptor subtypes coupled to inositol lipid metabolism, M1 and M4, can elicit a clearcut mitogenic response, whereas two others, M2 and M3, which are coupled to negative regulation of adenylate cyclase, are relatively inefficient in eliciting DNA synthesis. (The nomenclature in this field is confusing — see Eric Barnard's News and Views article². The receptors designated here¹ M3 and M4 are referred to by many as M4 and M3, respectively.)

Capon and colleagues extend these observations to a physiologically more relevant preparation — enriched primary cultures of astrocytes from the developing brain — and find that subpopulations of astrocytes show a mitogenic response to carbachol. Thus they suggest that acetylcholine, in addition to its transmitter function, may act as a neuron-to-glia cell signal in triggering cell division. These results are part of a growing consensus that not just growth factors but a large variety of chemical messengers are able to drive DNA synthesis in the mature and the developing nervous system.

It now seems that several neural receptors coupled to phospholipase C hydrolysis of inositol lipids have mitogenic potential, including the cloned 5-HT_{1c} (ref. 3) and *mas* oncogene/angiotensin receptors⁴. All the muscarinic receptors so far identified can drive DNA synthesis to varying extents, implying that receptors coupled

to different signalling pathways, including the negative regulation of adenyl cyclase through the GTP-binding protein G_i (ref. 5), can drive the common output event of cell proliferation. Indeed, it could well turn out that all the main second-messenger pathways regulated by G-protein-coupled receptors are conditionally mitogenic, albeit to varying extents.

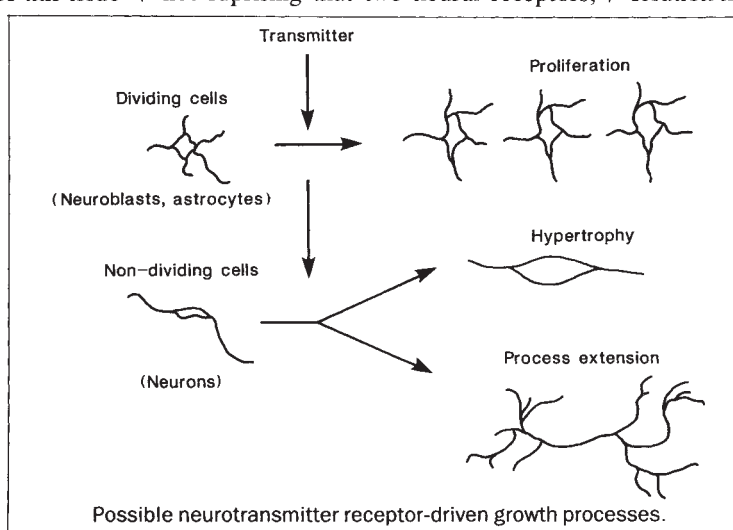
Against this background, it is perhaps not surprising that two neural receptors,

control (see figure). The neuronal counterpart of a hypertrophic response would not be the induction of increased somatic volume, but rather would be the stimulation of increased cytoplasmic surface area by enhancing process formation and branching. This type of 'trophic' activity is not normally considered to be within the repertoire of neurotransmitter action, but the notion that trophic forms of growth responses lie on one end of a continuum which has cell division at the other implies an appealing genetic unity to all forms of neural growth.

Neurotransmitter stimulation of proliferation must also be dependent on some

other cellular events which help to define the permissive 'context'. In *mas*-transfected neural cells⁴, angiotensin acts as a direct mitogen but bradykinin does not, even though both receptors are coupled to inositol-lipid hydrolysis. There must therefore be some aspect of the angiotensin-mediated activation of these cells which is essential for a mitogenic response.

One candidate for an essential element in mitogenic action which is being widely examined is the newly defined lipid kinase called phosphatidylinositol 3-kinase or type I



mas and the 5-HT_{1c} receptor, can induce a malignant state of uncontrolled growth, leading to tumours when expressed in NIH-3T3 cells^{4,6} — the standard indicator for oncogene identification. Thus, some neurotransmitter receptors may not only have the unexpected characteristic of mitogenic control, but also can be properly considered 'proto-oncogenes'. There is a need to assess the entire family of cloned G-protein-coupled receptors, which share the seven-hydrophobic domain pattern in amino-acid sequence⁷ for transforming or tumorigenic activity in susceptible cells.

The action of mitogenic neurotransmitter receptors is most readily understood in the context of early nerve cell precursors, which are still actively dividing, or in the context of non-neuronal cells such as astrocytes, which maintain a proliferative capacity in adulthood. The possibility that mitogenic neurotransmitter receptors act in novel ways on post-mitotic neurons should also be considered, however.

In cardiac myocytes, another non-dividing cell population, many of the genetic responses associated with cell proliferation are maintained in receptor-stimulated hypertrophy, where a cell does not divide but gains cytoplasmic mass⁸. The genetic cascade, which in immature neuroblasts is exclusively linked to proliferation, may therefore be retained in mature neurons for other forms of growth

PI kinase, which makes an unusual inositol phospholipid, phosphatidylinositol 3-phosphate⁹. The reason for the intense interest in this enzyme is that it seems to associate directly with the mitogenic platelet-derived growth factor receptor through a domain inserted in the tyrosine kinase, and the deletion of this domain leads to an abolition of mitogenic responses through the receptor¹⁰. Other evidence suggests that type I PI kinase also forms active complexes with the viral-transforming gene polyoma middle-T antigen¹¹. Although it is not at all clear how this enzyme acts to define the mitogenic context, it provides a provocative candidate 'mitogenic switch' which might act in nerve cells, as in other cell populations, to specify whether biochemical transduction of a neurotransmitter signal leads to a specific growth response. □

Michael R. Hanley is in the MRC Molecular Neurobiology Unit, University of Cambridge Medical School, Cambridge CB2 2QH, UK.

1. Ashkenazi, A., Ramachandran, J. & Capon, D.J. *Nature* **340**, 146–150 (1989).
2. Barnard, E. *Nature* **335**, 301 (1988).
3. Julius, D. et al. *Science* **244**, 1057–1062 (1989).
4. Jackson, T.R. et al. *Nature* **335**, 437–440 (1988).
5. Seuwen, K. et al. *Nature* **335**, 254–256 (1988).
6. Young, D. et al. *Cell* **45**, 711–719 (1986).
7. Dohlman, H.G. et al. *Biochemistry* **26**, 2657 (1987).
8. Simpson, P.C. A. *Rev. Physiol.* **51**, 189–202 (1988).
9. Whitman, M. et al. *Nature* **332**, 644–646 (1988).
10. Coughlin, S.R. et al. *Science* **243**, 1191–1194 (1989).
11. Courtneidge, S.A. & Heber, A. *Cell* **50**, 1031–1037 (1987).

Magnetism of layered helium-3

P. V. E. McClintock

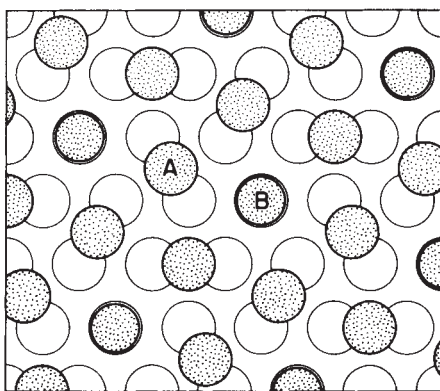
A LAYER of two of ^3He atoms adsorbed on a flat surface at very low temperatures constitutes an almost ideal two-dimensional system in which a whole range of thermal and magnetic quantum phenomena may conveniently be investigated. Recent experiments, in several laboratories, have revealed some strange and unexpected behaviour that can arise in an apparently incomplete second atomic layer. Veit Elser has risen to the challenge provided by these data and, writing in *Physical Review Letters* (62, 2405–2408; 1989), suggests that the observed effects are associated with nuclear antiferromagnetism in what is actually a peculiarly loosely knit kind of two-dimensional solid.

The experiments that Elser discusses were performed using ^3He adsorbed on Grafoil, a form of graphite that has been 'exfoliated' so as to separate the planes of atoms, thus creating an enormous effective surface area, typically of 50 m^2 within a volume of only 1 cm^3 . The surface thus created is atomically flat to an excellent approximation, albeit broken into sections approximately $1\text{ }\mu\text{m}$ across. A helium atom adsorbed onto such a surface enters an almost ideal two-dimensional world in which it can move freely in any direction parallel to the surface but where, at low temperatures, motion perpendicular to the surface, or escape from it, is prevented by the interatomic forces. The two-dimensional analogues of many familiar three-dimensional phenomena can be reproduced in this system. As the temperature and the surface density of adsorbed helium atoms (the coverage) are varied, it is possible to identify two-dimensional gases, liquids and — at high enough packing densities — crystalline solids.

The question at issue relates to what happens when a second layer of ^3He is adsorbed on top of an already completed, highly compressed, tightly packed first layer. The latter is known to form a triangular lattice that is incommensurate with the underlying graphite: that is, the periodicities of the two lattices are unrelated. The second layer is potentially of particular interest because, being further away from and, for that reason, less tightly bound to the underlying graphite, it can be much looser and more readily able to reflect subtle influences such as quantum-mechanical exchange interactions between the ^3He atomic nuclei, which carry tiny magnetic moments. Experiments to investigate weak interactions of this kind must inevitably be done at ultra-low temperatures in order to reduce the otherwise dominant

influence of thermal fluctuations.

The microscopic theory of exchange interactions is in general extremely difficult. In practice it is hard to calculate even the sign of the interaction for any given arrangement and spacing of atoms, and thus to predict in advance whether a given structure will be ferromagnetic (parallel magnetic moments on adjacent sites) or antiferromagnetic (antiparallel moments). In fact, it is more usual to start from the experimental results, and then to work backwards to show that the observations can be explained by a suitable choice of exchange constants. In the case of ^3He on



Sketch to illustrate the structure proposed for the atoms (shaded) of the second layer of ^3He . Both of the layers are two-dimensional solids, and they both have triangular lattices, but the solid that is formed by the second-layer atoms exists at much the lower density. From V. Elser (*Phys. Rev. Lett.* 62, 2405–2408; 1989).

Grafoil, magnetization measurements by H. Franco, R.E. Rapp and H. Godfrin (*Phys. Rev. Lett.* 57, 1161–1164; 1986) demonstrated the existence of a strong maximum at a surface coverage of $0.24\text{ atoms }\text{\AA}^{-2}$, corresponding to completion of a nuclear ferromagnetic solid second layer. The authors also observed a much weaker — and quite unexpected — magnetization peak which occurred at a coverage of only $0.18\text{ atoms }\text{\AA}^{-2}$. It is this weaker peak that forms the subject of Veit Elser's paper.

The magnetization peak in question is puzzling because it is seen at a coverage where the second layer should apparently be fluid and therefore, presumably, non-magnetic: that is, it occurs well below the $0.24\text{ atoms }\text{\AA}^{-2}$ corresponding to completion of the second layer, but considerably above the $0.114\text{ atoms }\text{\AA}^{-2}$ at which the first layer is completed.

There are three other significant pieces of experimental information that must be taken into account. First, earlier heat-capacity measurements by S. W. Van Sciver and O. E. Vilches (*Phys. Rev. B* 18,

285–292; 1978) gave evidence of a melting peak close to the relevant $0.18\text{ atoms }\text{\AA}^{-2}$ coverage, thus implying that this is a density at which the second layer should be solid at low temperatures. Second, new high-precision heat-capacity measurements by D. S. Greywall and P. A. Busch at a similar coverage (*Phys. Rev. Lett.* 62, 1868–1871; 1989) not only confirm the data of Van Sciver and Vilches, but also reveal the existence of a sharp peak at a temperature of about 2.5 mK . And third, in apparent contradiction, a neutron-diffraction experiment by C. Tibby, H. Wiechert, H. J. Lauter and H. Godfrin (*Physica B & C* 107, 209–210; 1981) had apparently shown that the second layer at a coverage of $0.18\text{ atoms }\text{\AA}^{-2}$ is not in fact solid at all.

Elser's proposed solution to the conundrum lies in the two-dimensional structure sketched in the figure. His suggestion, developing an idea of Greywall and Busch, is that the second-layer (shaded circles) at the coverage in question is indeed a solid, albeit a low-density one, that is 'registered' on (is commensurate with) the first layer. The second-layer atoms are separated from the underlying graphite by a distance such that they must be virtually unaffected by the positions of individual graphite atoms and hence will have a negligible tendency to form a lattice commensurate with the graphite lattice. Rather, they will be dominated by the positions of the first-layer atoms. Even at the low second-layer density in question, interactions between the layers can give rise to a registration effect and, in order to correspond to the measured surface coverage, Elser suggests the rather curious form shown in the figure. There are two different kinds of second-layer sites: A-sites in which the atom sits in the local potential minimum (actually a saddle) between a pair of neighbouring first-layer atoms; and B-sites in which a second-layer atom lies at the potential maximum immediately above a first-layer atom. There are three A-sites to every B-site, so the arrangement is energetically favourable.

The suggested structure, with a registered two-dimensional solid existing even at the low second-layer coverage in question, immediately provides a qualitative explanation for the heat-capacity and magnetization data. There remains, of course, the vexing question of why the neutron diffraction experiment failed to find any evidence for a solid second layer. However, the looseness of the structure, combined with the probable importance of near-neighbour hard-core repulsion in stabilizing it, implies rather large quantum mechanical zero-point fluctuations of the atoms about their equilibrium positions: this, Elser suggests, is the most likely explanation of the negative neutron-diffraction result. He goes on to explore

in detail the magnetic and thermal properties to be expected of such a system and he demonstrates plausibly that they are consistent with the experimental observations made to date. It seems that the specific heat anomaly at 2.5 mK is associated with a partial antiferromagnetic transition in which only the A-site atoms participate.

The low-density second layer of ^3He

has some unusual properties and clearly merits further investigation, not least because Elser's model also predicts that an additional phase transition should occur at even lower temperatures, perhaps near 0.4 mK. □

P. V. E. McClintock is in the Department of Physics, University of Lancaster, Lancaster LA2 4YB, UK.

CENTROSOMES

Organizing cytoplasmic events

Douglas R. Kellogg

ALTHOUGH the existence of the centrosome has been known for more than 100 years, it remains enigmatic and poorly understood. Modern studies have identified it as a structurally complex organelle responsible for the nucleation of microtubules^{1,2}. Although it is clear that centrosomes function as the two poles of the mitotic spindle responsible for the ordered segregation of chromosomes during cell division, the full range of their activities remains to be discovered. Raff and Glover³ have recently reported the use of the DNA-synthesis inhibitor aphidicolin to uncouple centrosomal and nuclear behaviour in the early *Drosophila* embryo. These experiments dramatically demonstrate the importance of the centrosome as an organizer of cytoplasmic events.

During normal development of the early *Drosophila* embryo⁴, the first 14 nuclear divisions take place in a syncytium. The nuclei initially divide in the interior of the embryo, but after the eighth division most of the nuclei begin to migrate to the embryo cortex. A small group of nuclei at the posterior end of the embryo are the first to reach the cortex and, after dividing once, become incorporated into cells by a pinching off of the plasma membrane. These 'pole cells' constitute the future germ line. The rest of the nuclei reach the embryo cortex at the beginning of the tenth nuclear cycle, and form an even monolayer over the surface of the embryo. They go through four more synchronous divisions and then become cellularized by synchronous invaginations of the plasma membrane during interphase of the fourteenth nuclear cycle. The movements of gastrulation begin almost immediately thereafter.

Each of the nuclei in the syncytial *Drosophila* embryo is surrounded by its own unique domain of cytoplasm^{5,6}. This can be seen in the light microscope as a cleared region surrounding each nucleus that excludes yolk granules and other large particles. Immunocytochemical staining shows that each of these domains is defined by a system of cytoskeletal fibres. When the nuclei are at the surface of the embryo, actin filaments form a

meshwork between each interphase nucleus and the plasma membrane (an 'actin cap'), while microtubules are organized in each domain by a pair of centrosomes in close association with the nuclear envelope. These two types of cytoskeletal elements undergo characteristic changes in their distribution during each cell cycle. The results of Raff and Glover³ suggest that it is the centrosome, and not the nucleus, that organizes these cytoplasmic events.

Raff and Glover show that if embryos are injected with aphidicolin at approximately the ninth nuclear cycle, when the nuclei have largely completed their migration to the embryo cortex, further nuclear division is blocked⁷ but the centrosomes become dissociated from the blocked nuclei and continue to divide. The authors were not able to observe the duplication of centrosomes in real time, but they infer from the appearance of fixed embryos that the centrosomes duplicate synchronously, and that the duplication probably takes

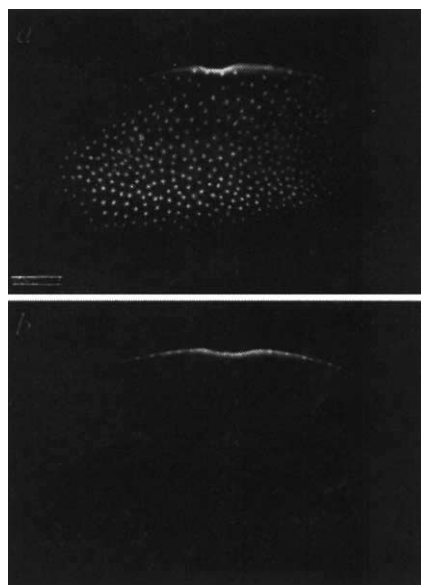
place in the expected 1→2→4→8 fashion.

If aphidicolin is injected during the eighth nuclear cycle, when the nuclei are just beginning their migration to the embryo cortex, both nuclear division and nuclear migration are blocked. Under these conditions, the centrosomes in 65 per cent of the embryos continue to divide and migrate to the embryo cortex (see figure). Remarkably, the centrosomes that arrive at the posterior end of the embryo become incorporated into pole cells that contain no nuclei. Each pole cell contains one centrosome or centrosome pair, and observation of the process of pole-cell formation by phase-contrast microscopy shows that it is indistinguishable in aphidicolin-injected and in control embryos.

Several observations indicate that the formation of these pole cells requires the presence of centrosomes. First, in the 35 per cent of the embryos injected with aphidicolin, in which the centrosomes do not migrate to the embryo cortex, no pole cells are formed. In addition, if embryos are injected with colchicine to depolymerize microtubules at the eighth nuclear cycle, both nuclear and centrosomal migration are blocked, and no pole-cell formation is initiated. These observations strongly suggest that it is the centrosomes that direct pole-cell formation, presumably mediated by microtubules.

In addition to forming pole cells, the centrosomes that can migrate to the embryo cortex in aphidicolin-treated embryos can organize dynamic cytoplasmic domains resembling those surrounding nuclei in untreated embryos. Immunocytochemical staining shows that each centrosome at the embryo cortex is overlain with an actin cap that closely resembles the actin caps seen over nuclei in untreated embryos.

Similar experiments were carried out previously by Sluder *et al.*⁸ using sea urchin embryos, and by Nagano *et al.*⁹ and Picard *et al.*¹⁰ using starfish embryos. These authors dissociated nuclear and centrosomal behaviour either by treatment with aphidicolin or by removal of the nucleus with a micropipette. Using Nomarski optics, the behaviour of centrosomes can be observed in living embryos by virtue of the astral arrays of microtubules that they



Distribution of centrosomes in an aphidicolin-injected *Drosophila* embryo. Focusing at the surface (a) and interior (b) of the embryo shows that centrosomes have migrated to the surface. Bar = 50 μm . (From ref. 3, courtesy of J. Raff.)

1. Karsenti, E. & Maro, B. *Trends biochem. Sci.* **11**, 460–463 (1986).
2. Vorobjev, I.A. & Nadezhdina, E.S. *Int. Rev. Cytol.* **106**, 227–291 (1987).
3. Raff, J.W. & Glover, D.M. *Cell* **57**, 611–619 (1989).
4. Foe, V.E. & Alberts, B.M. *J. Cell Sci.* **61**, 31–70 (1983).
5. Karr, T.L. & Alberts, B.M. *J. Cell Biol.* **102**, 1494–1509 (1986).
6. Warn, R.M., Magrath, R. & Webb, S. *J. Cell Biol.* **98**, 156–162 (1984).
7. Raff, J.W. & Glover, D.M. *J. Cell Biol.* **107**, 2009–2019 (1988).
8. Sluder, G., Miller, F.J. & Rieder, C.L. *J. Cell Biol.* **103**, 1873–1881 (1986).
9. Nagano, H., Hirai, S., Okano, K. & Ikogoni, S. *Dev. Biol.* **85**, 407–415 (1981).
10. Picard, A., Harricane, M.-C., Labbe, J.-C. & Doree, M. *Dev. Biol.* **128**, 121–128 (1988).

nucleate. Sluder *et al.*⁸ find that the centrosomes in enucleated sea urchin embryos undergo division in a precise 1→2→4→8 fashion, with a slightly slower cycle time than control embryos. Electron microscopy reveals that these centrosomes contain the normal complement of centrioles, suggesting that the centrosomes undergo normal division rather than fragmentation. Although the enucleated sea urchin embryos do not undergo cytokinesis, cleavage furrows are initiated to a variable degree. By contrast, starfish embryos treated with aphidicolin can undergo many rounds of cytokinesis.

It is worth considering the role of the centrosome as a cytoplasmic organizer in a developmental context. The nuclei that migrate to the posterior pole of the

Drosophila embryo must interact with specifically localized determinants that ultimately cause them to become pole cells. The finding that pole cells can form in the absence of nuclei suggests that some of these determinants become incorporated specifically into the pole-cell lineage by virtue of an interaction with the centrosome or its associated microtubule networks. In any case, it is clear that an understanding of the spatial and temporal organization of the early *Drosophila* embryo cytoplasm is integral to our understanding of the process of pattern formation. □

Douglas R. Kellogg is in the Department of Biochemistry and Biophysics, University of California School of Medicine, San Francisco, California 94143-0448, USA.

MATERIALS SCIENCE

Less thick than a brick

Paul Calvert

INTELLIGENT materials seem to have become a new hot topic: the US army is about to inaugurate a research initiative to develop them; and recent meetings in Japan and the United States indicate dawning interest amongst materials scientists. But although several specific examples were presented at the latest meeting*, it is still not obvious exactly what constitutes an intelligent material. Nevertheless, the concept is attractive and clearly leads in the right direction for the achievement of a higher level of integration between machines and the materials used to construct them.

The hydrogels are one group of materials that have some of the right characteristics. These are water-swollen crosslinked polymers, usually modified acrylates or acrylamides. The physics of these systems has been studied in detail by Tanaka and co-workers¹ and other groups are working on devices². If the gel contains anionic groups such as carboxylic acids, the equilibrium swelling is very sensitive to the ionic strength of the surrounding liquid. By moving the gel from a salt solution to water, and by applying an electric field to drive out the counter ion, the gel can be made to contract. The action is superficially similar to that of muscle. This ability of a material to respond to a change in its surroundings must be part of the concept of intelligence. In practical terms, the response is slow (taking 1–10 s, much more than the 10–100 ms taken by muscle fibres) and the muscle action is rather weak (1 mW cm⁻²; muscle delivers 1 W cm⁻²). But work is continuing on devices for robotics, for controlled drug delivery and for variable porosity membranes.

Craig Rogers (Virginia Polytechnic) has

been working on 'smart systems' which include the ability to damp vibrations actively. A framework of struts with strain gauges and motors that can change the length of a strut, can be programmed to counteract vibrations of a particular frequency. It tackles a problem similar to that of walking without spilling a cup of coffee. This work has now been extended to studies of composite materials containing fibres of shape-memory alloy. Shape-memory alloys undergo large, reversible changes of shape when heated through a transition temperature. By selective electrical heating of some of the fibres, the stiffness and vibrational modes of the beam can be tuned³.

Danilo de Rossi (Università degli Studi di Pisa) has been working on an artificial skin which could be used for robots. Such a material should be able to sense changes in temperature and heat flow, such as allow us to distinguish thermal conductivity when we touch cold metal or plastic. It should also be able to sense the full six components of stress, not just the pressure perpendicular to the skin, and have good spatial resolution. In this way one would hope to be able to distinguish surface characteristics such as roughness, friction and stickiness. In skin the response may be associated with the piezoelectric behaviour of keratin plus a pressure sensitivity in the underlying dermis.

Rossi has been investigating various types of sensors, including a pseudo-epidermis of multiple layers of piezoelectric polymers such as polyvinylidene fluoride. The dermis is modelled by a layer of fibre-reinforced polyelectrolyte gel with embedded electrodes acting as sensors for pressure and strain rate. This material shows a further aspect that one might associate with intelligence,

the ability to sense changes in the surroundings using multiple channels of information⁴.

A hollow composite beam can be filled with electro-rheological fluid which becomes rigid under an applied field and so make the beam stiffer (M.V. Ghandi, Michigan State University). Coupled with sensors, this would allow the beam to change properties in response to changes in load. Piezoelectric materials acquire a voltage when strained, and contract, expand or bend when charged electrically. They are thus an obvious starting point for assembling an intelligent structure. R. E. Newnam (Pennsylvania State University) is working on combinations of piezoelectric sensors and actuators which will contract away from a surface when they touch it, so giving the hard material an apparently soft touch.

High-performance composite materials for aircraft are mainly carbon-fibre-reinforced polymer. By incorporating a few silica fibres one can use spectroscopic techniques to probe the degree of cure, degradation or stress levels inside the material. In principle this information could be used at least to make the material declare its imminent failure, if not to initiate its own repair.

From these studies one can begin to see what should be the properties of an intelligent material. It should be able to sense changes in its environment in some detail. It should contain some actuator function to allow a response to those changes. The link should be governed by a logic system although the 'brain' is not necessarily embedded in the material. Obvious models can be found in biological materials and several were discussed at the meeting. One characteristic is the ability to restructure in response to external stress, as bone remodels when continually loaded non-uniformly. Also, intelligent materials should be able to remodel in the face of high local stresses, corrosion or developing fatigue. This means that we should start to think of materials not necessarily as closed systems, but as being able to change composition or to grow.

Whether 'intelligent' or 'smart' are the right descriptions is an open question, as is the appropriateness of the term 'materials'. The last word should probably go to Humpty Dumpty⁵: "When I use a word, it means just what I choose it to mean — neither more nor less". □

Paul Calvert is in the Arizona Materials Laboratory, University of Arizona, Tucson, Arizona 85712, USA.

* International Workshop on Intelligent Materials, Tsukuba, 15–17 March 1989.

1. Tanaka, T. *et al.* *Phys. Rev. Lett.* **45**, 1636–1639 (1980).
2. Osada, Y., Kishi, R. & Hasebe, M. *J. Polymer Sci. Lett.* **25**, 481–485 (1987).
3. Rogers, C.A. (ed.) *Smart Materials, Structures and Mathematical Issues* (Technomic, 1989).
4. DeRossi, D., Domenici, C. & Chiarelli, P. in *Sensors and Sensory Systems for Advanced Robots* (ed. Dario, P.) (NATO ASI Series, Springer, Berlin, 1988).
5. Carroll, L. *Through the Looking Glass* (Macmillan, London, 1956).

Goslings of gay geese

Jared M. Diamond

HOMOSEXUAL female pairs of wild birds — already reported for four gull, one tern and one goose species — have aroused much interest¹⁻⁷. Does the phenomenon reflect a hormonal or behavioural idiosyncrasy of the homosexual females that distinguishes them from heterosexual females, one that produces no evolutionary benefit (no offspring)? Or is it instead an adaptation compatible with reproduction and favoured by natural selection under conditions of male scarcity? A recent study by Quinn *et al.*¹ demonstrates the feasibility of the latter interpretation.

Recognition of the phenomenon stemmed from the observation that up to 14 per cent of western gull clutches on Santa Barbara Island, California, contained twice as many eggs as do normal clutches². Although this observation might in principle have arisen from other causes (such as polygynous females sharing a nest, brood parasitism or egg dumping), the prevalent cause turned out to be pairs consisting of two females, both of whom laid eggs. Like male/female pairs of gulls, paired females often make lasting bonds and nest with each other in successive years in the same colony, even at the same nest site^{2,3}. Homosexual female pairs of western gulls practise the same courtship and territorial behaviour as do heterosexual pairs, except that courtship feeding is rare and only three homosexual pairs

exhibited attempted copulation².

Homosexual pairing is often associated with a shortage of males in the nesting colony^{4,5}. The ratio of males to females collected since 1950 among adult western and herring gulls is only 0.42–0.43, for example, because of selective post-fledging mortality of males⁵. It is interesting that above-average sized clutches have been

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Natural herstory — the lesser snow goose. (Courtesy of Dr Fred Cooke.)

regularly observed in these two species only since 1950; before 1950, the sex ratio was closer to 1.0. Experimental removal of many males from California and ring-billed gull colonies resulted in a threefold increase in frequency of 'supernormal' clutches compared with control colonies⁶.

For homosexual pairing in gulls to be explicable by natural selection at the individual level, the female must somehow become fertilized and succeed in rearing offspring. Western gull female pairs provided experimentally with fertile eggs from other nests do incubate, feed and fledge young with virtually normal success rates². Among the eggs found naturally in supernormal clutches, many are infertile but some are fertile. How do paired females become fertilized?

Quinn *et al.*¹ now report results of a restriction-fragment-length polymorphism analysis to determine parentage of goslings hatched by a female pair of lesser snow geese on Hudson Bay, Canada. When discovered, the nest contained eight eggs, twice the normal clutch size for the species. Seven of the eggs hatched. Both females defended the goslings, and removal of one female led the other to give loud 'mate calls' typical of male/female pairs. The

genetic analysis indicates that goslings 1, 2 and 3 stemmed from one of the females fertilized by one male; goslings 4, 6 and 7 stemmed from the other female fertilized by a different male; and gosling 5 could have stemmed from either female but had to have been sired by yet a third male. Thus, both females had become fertilized independently and were not the co-widows of one deceased or absent polygynous male.

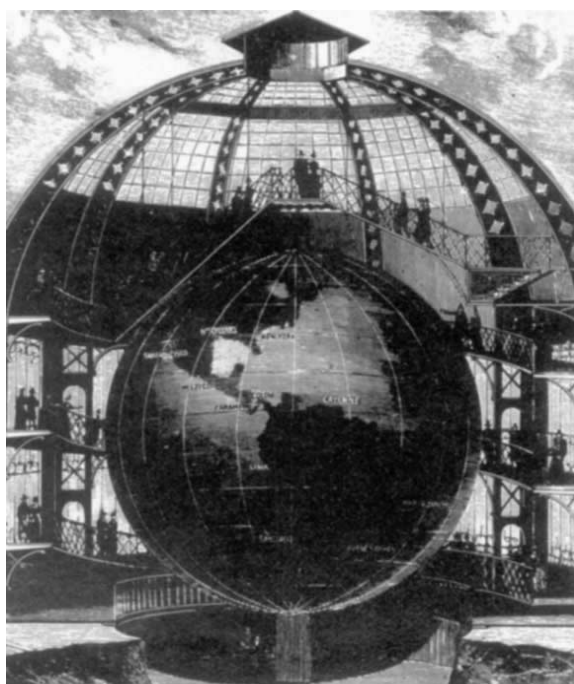
This is the first observation of female pairing in lesser snow geese, a species for which the sex ratio is not known to be skewed. Although the advantage of female pairing is thus uncertain for the geese, its adaptive value now becomes clear for the gull colonies having a shortage of males. One gull alone is incapable of rearing young because other gulls destroy the clutch when the single parent leaves to feed. Normal nests of male/female pairs are always attended by one parent. In a nominally monogamous colony a male deficit means that some females cannot obtain male mates. But any female can still become fertilized (at least in gulls) by some nominally monogamous male, as mated males seek to copulate with females other than their mate. Once fertilized, two females can rear the young almost as well as one female plus one male.

These observations raise the perennial question of why males exist at all at a sex ratio near 1.0. After a male gull has contributed semen, he appears to play almost no role that a female cannot play equally well. It is true that female gulls do not provide each other with the courtship feeding that male gulls provide their mates, with the result that eggs of homosexual female pairs are smaller and may have poorer posthatching survival than do eggs of heterosexually paired females². However, if long-term reproductive success per egg is at least 50 per cent of normal, homosexual pairs would still have a higher reproductive output per individual than do heterosexual pairs.

It is also true that, in species whose males are much bigger than females (unlike male gulls), males are useful for protecting the young. Yet other males themselves are one of the main threats in the first place. Further study of homosexually paired female birds may help clarify what, if anything, males are good for — in an evolutionary sense, of course. □

Jared M. Diamond is a professor of Physiology at the University of California Medical School, Los Angeles, California 90024–1751, USA.

100 years ago From *Nature* 40, 280; 18 July 1889.



One of the great attractions of the Paris Exhibition is a great terrestrial globe, one millionth of the actual size of the earth.

1. Quinn, T.W. *et al.* *Auk* **106**, 177–184 (1989).
2. Hunt, G.L. & Hunt, M.W. *Science* **196**, 1466–1467 (1977).
3. Kovacs, K.M. & Ryder, J.P. *Auk* **98**, 625–627 (1981).
4. Hunt, G.L. *et al.* *Auk* **97**, 473–479 (1980).
5. Conover, M.R. & Hunt, G.L. *Wilson Bull.* **96**, 619–625 (1984).
6. Conover, M.R. & Hunt, G.L. *Condor* **86**, 472–476 (1984).
7. Conover, M.R. *Condor* **86**, 467–471 (1984).

A case of two conductors

T. J. Shankland

ELECTRICAL conductivity of the lower continental crust has presented a persistently irksome problem for interpretation in geological terms. From geophysical techniques such as magneto-tellurics, the lower crust seems to be a fairly good electrical conductor, having conductivities above 10^{-6} S m^{-1} , usually in the range 10^{-4} – 10^{-2} S m^{-1} ; yet laboratory measurements on typical crustal rocks seem to fall orders of magnitude below these values unless temperatures approach the melting point¹. We are driven to hypothesize that *in situ* crustal rocks contain other conducting phases that seem absent in most specimens used for laboratory measurements. On page 136 of this issue² Bailey *et al.* argue that the elusive conductive phases are saline fluids, whereas on page 134 Frost *et al.*³ present a case for the presence of graphite precipitated from a CO_2 -rich fluid during cooling. Because sea water and silicate melts have conductivities of the order of 3 S m^{-1} and graphite, 10^4 S m^{-1} , only small quantities of either material are required to explain the small conductivities above, provided that the conducting phases are interconnected.

These alternative hypotheses may not be as mutually exclusive as they appear. For one thing, Bailey *et al.* describe field measurements in a region for which there are not samples obtained from depth, whereas Frost *et al.* investigate a conductive phase in metamorphic rocks that they consider representative of the lower crust and that were obtained at the surface in a region where crustal conductivity was not measured. In both cases the arguments could apply to old shield areas (tectonically stable regions) where temperatures are relatively low so that partial melt is not an issue. (In active metamorphic zones or volcanic regions one would expect fluids and melts which are near lithostatic pressure and are capable of broad crustal permeation.)

Bailey *et al.* find horizontal conducting layers that cross a tilted Proterozoic thrust fault, the Ivanhoe Lake cataclastic zone in northern Ontario, and attribute the layers to the presence of conductive fluids. In contrast to the observed layered structure, if the conducting phases were differing contents of sulphides or graphite in the structures on either side of the fault, then it would be reasonable to expect differing conductivities at different depths on opposite sides of the fault. However, there is only a slight break where the fault intersects the conducting layers. There is a substantial history of quantitative interpretation of crustal conductivity in terms of fluid content^{4,5}, and superdeep holes in ancient crust⁶ reveal surprising amounts of

fluid, consistent with at least the conducting layers in the upper 6 km observed by Bailey *et al.* But there remains some petrological discomfort⁷ with the notion that interconnection of a conducting fluid phase could persist over geological times without its incorporation into relatively insulating hydrated phases⁸ at warmer temperatures below 10–15 km, the depth of their deeper high conductivity layer, unless the fluid is sealed by hydrated minerals in fractures rather than on a grain boundary scale¹.

Frost *et al.* use a scanning auger multi-probe to identify grain boundary coatings of carbon having thickness of the order of 1,000 Å; mapping this thin layer is a remarkable advance in petrographic observation. The authors' calculation of electrical conductivity for an intergranular conducting fabric suggests that a conducting phase having this texture could account for observed crustal conductivities. Such an approach also has other support. Mathez and Delaney⁹ observed grain-boundary carbon on intergranular surfaces in mantle olivine and favoured a similar, late-stage carbon precipitation mechanism rather than a permanent mantle condition. Sternberg¹⁰ actually described high carbon contents in rocks from the Canadian shield at the same location where he measured conductivity profiles.

These two papers break new ground in their different kinds of observation. To the extent that their work can be taken as representative of other geological regions, they have narrowed the range of speculation about solutions of a long-standing problem to two mechanisms. With more work along these lines we may find ourselves looking not for a single explanation that works in nearly every case but for the explanation that fits each region. □

T. J. Shankland is in the Geophysics Group and the Institute of Geophysics and Planetary Physics, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA.

- Shankland, T.J. & Ander, M.E. *J. geophys. Res.* **88**, 9745–9784 (1983).
- Bailey, R.C., Craven, J.A., Macnae, J.C. & Polzer, B.D. *Nature* **340**, 136–138 (1989).
- Frost, B.R., Fyfe, W.S., Tazaki, K. & Chan, T. *Nature* **340**, 134–136 (1989).
- Brace, W.F. in *The Structure and Physical Properties of the Earth's Crust*, Geophys. Monogr. Ser. vol. 14 (ed. Heacock, J.G.) 243–255 (Am. geophys. Un., Washington DC, 1968).
- Hyndman, R.D. & Hyndman, D.W. *Earth planet. Sci. Lett.* **4**, 427–432 (1968).
- Lavrusenko, I.P., Byeleviski, G.A., Prizov, Y.Y. & Stolyarov, D.E. *The Superdeep Well of the Kola Peninsula* (Springer, Berlin, 1986).
- Wannamaker, P.E. *J. geophys. Res.* **91**, 6321–6327 (1986).
- Olhoef, G.R. *J. geophys. Res.* **86**, 931–936 (1981).
- Mathez, E.A. & Delaney, J.R. *Earth planet. Sci. Lett.* **56**, 217–232 (1981).
- Sternberg, B.K. *J. geophys. Res.* **84**, 212 (1979).

Against the law

THESE days, technical progress is being increasingly obstructed by legal process. In the United States, that cutting edge of capitalism, some products like vaccines, anaesthetics and light aircraft have been almost driven off the market by the risk of product-liability lawsuits. Others like step ladders and medical services have prices heavily inflated by insurance against legal claims. Daedalus, whose product liability problems at DREADCO are clearly extreme, has in self defence set his mathematicians to develop a computer model of the legal process.

Their basic assumption is that the rights and wrongs of the matter can be ignored. The chance of a lawsuit succeeding is a function purely of the money deployed by either side: the fees of the retained lawyers (or in the case of 'no win, no fee' suits, their income from past wins) and the insurance premiums paid against the liability. All these sums depend on the outcome of past actions, and therefore present a good numerical guide to this one. The likely award will also be a function of time; for award levels seem to increase monotonically. The various parameters of SUE (as the model is called) will be adjusted till it correctly predicts the outcome of most past lawsuits. It will then enable DREADCO to sue others, or react to being sued itself, by pure game theory, without the need for costly legal advice.

With the usual naive logic of computer routines, SUE has already generated some intriguing strategies. One is a 'negative insurance policy' in which the insurance company pays premiums to DREADCO in return for a huge payout to the insurers if DREADCO is sued. Legal action will therefore instantly bankrupt DREADCO, leaving the plaintiff with no prospect of any profit. This 'poison-pill' should be a powerful deterrent.

But when SUE has fully proved itself as a robust and effective program, Daedalus plans to sell it in the open market. Once every potential litigant has SUE and knows that his opponent has it as well, litigation will lose much of its point. All concerned will be able to simulate their legal action instead of having to fight it out for real. Not being egged on by expensive and ingenious lawyers, they will probably soon reach an amicable settlement.

The legal profession will doubtless fight back. It will attempt to outflank SUE by ever wider and more generous judgements of liability. The logical end of this process, muses Daedalus, would be the final assertion of everybody's limitless liability to everybody else for everything. So just when the socialist world is grudgingly seeing the virtues of capitalism, the capitalist world could find itself legally obliged to adopt the purest socialism.

David Jones

Leucine-zipper motif update

SIR—The leucine-zipper motif, consisting of four to five leucines repeated every seventh amino-acid residue, mediates dimerization of some DNA-binding proteins¹⁻⁴ but has also been found in several other proteins, sometimes adjacent to proposed transmembrane regions⁵. Thus, as has been previously suggested for coiled-coil structures⁶, many protein-protein interactions may be mediated through leucine zippers, which may actually constitute a subset of the coiled-coil motif⁷. We now report the presence of a leucine zipper in each of a set of sequences thought to represent a family of voltage-gated K⁺ channels and suggest its involvement in subunit interactions⁸ that may mediate voltage-dependent opening and closing of the channel.

The figure shows amino-acid sequence alignment of the voltage-gated K⁺ channels from the region containing the leucine-zipper motif. The leucine repeat is fully conserved except in the case of Shaw but the amino-acid substitutions in this case may not prevent subunit association¹⁴.

Interestingly, the leucine-zipper motif in these K⁺ channels is located immediately adjacent to the S4 domain. This domain is present in virtually all voltage-dependent channels and consists of the sequence Arg-X-X, repeated 4-8 times (where X is a hydrophobic amino acid and Arg can be replaced by Lys). It has been suggested that the S4 region forms an amphipathic transmembrane helix responsible for sensing membrane electrical potential⁹. In the favoured model of voltage-dependent gating, changes in membrane potential result in a translocation of the S4 helix¹⁰. But no mechanism has been proposed by which this translocation results in opening and closing of the channel.

The intimate association of the leucine-zipper motif with the S4 domain suggests that voltage-dependent translocation of the S4 sequence alters interactions of K⁺-channel subunits through the leucine zipper. This alteration may take place by physically pulling the zipper into or out of the membrane and/or bending it into or out of an α -helical configuration. This

would impart a voltage dependence to subunit interactions and provide a mechanism by which S4 translocation determines open and closed channel states. We are currently testing this hypothesis using site-directed mutagenesis.

We have also found leucine zippers within the proposed M4 transmembrane segments of at least some ligand-gated channels. Na⁺- and Ca²⁺-channel sequences also have leucine-zipper motifs next to their S4 domains. These motifs align with those in K⁺ channels but often contain prolines. It should be noted that Na⁺ and Ca²⁺ channels contain a large subunit consisting of 4 homology units, each of which is equivalent to a K⁺ channel subunit. Voltage-dependent control of homology domain association mediated by leucine zippers may also play a role in gating of Na⁺ and Ca²⁺ channels.

KEN McCORMACK
JAMES T. CAMPANELLI
MANI RAMASWAMI
MATHEW K. MATHEW
MARK A. TANOUYE

*Division of Biology,
California Institute of Technology,
Pasadena, California 91125, USA*

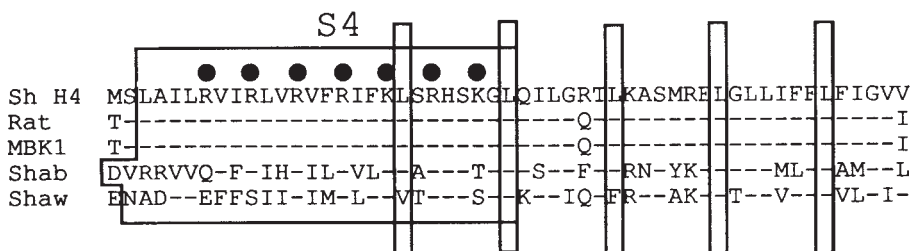
LINDA E. IVERSON

*Division of Neurosciences,
Beckman Research Institute
of the City of Hope,
Dart, California 91010, USA*

BERNARDO RUDY

*Department of Physiology and Biophysics,
New York University Medical Center,
New York, New York 10016, USA*

1. Kouzarides, T. & Ziff, E. *Nature* **336**, 646-651 (1988).
2. Landshultz, W.H., Johnson, P.F. & McKnight, S.L. *Science* **243**, 1681-1688 (1988).
3. Gentz, R., Rauscher, F.J., Abate, C. & Curran, T. *Science* **243**, 1695-1699 (1989).
4. Turner, R. & Tijan, R. *Science* **243**, 1689-1694 (1989).
5. Buckland, R. & Wild, F. *Nature* **338**, 547 (1989).
6. Cohen, C. & Parry, D.A.D. *Trends biochem. Sci.* **11**, 245-249 (1986).
7. O'Shea, E.K., Rutkowski, R. & Kim, P.S. *Science* **243**, 538-542 (1989).
8. Iverson, L.E. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5723-5727 (1988).
9. Guy, H.R. & Seetharamulu, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 508-512 (1986).
10. Catterall, W.A. *Science* **242**, 50-61 (1988).
11. Kamb, A. *et al. Neuron* **1**, 421-430 (1988).
12. Baumann, A. *et al. EMBO J.* **7**, 2457-2463 (1988).
13. Tempel, B.L. *et al. Nature* **332**, 837-839 (1988).
14. Butler, A. *et al. Science* **243**, 943-947 (1989).



Amino-acid alignment of K⁺ channels. Shaker (Sh)¹¹ and the rat¹² sequences have expressed functional channels in oocytes. The other sequences are from mouse (MBK1)¹³ and *Drosophila* (ref. 14). Amino acids identical to Sh are depicted by dashes. Leucines of the proposed zipper are boxed. The S4 region is also boxed, with circles indicating basic residues in Sh.

SIR—the leucine-zipper motif originally proposed for DNA-binding proteins¹ has recently been extended to the membrane F glycoproteins of paramyxoviruses and has been proposed to mediate their oligomerization². In this model, leucine or isoleucine residues are present at exactly seven-residue intervals for at least eight turns of an α -helix which facilitates dimerization with a second polypeptide with a similar motif.

We report the presence of this motif in the second membrane-spanning domain of several glucose-transporter glycoproteins. The glucose transporters of vertebrate cells consist of a family of at

	ref
CEF : tsLwslsvaIfsvggmIgsfsvsLff	12
Ftlsk: tsLwslsvaIfsvggmIgsfsvgLfv	9
HepG : ttLwslsvaIfsvggmIgsfsvgLfv	8
RatB : ttLwslsvaIfsvggmIgsfsvgLfv	10
IRGT : ttLwslsvaIfsvggmIgsfsvgLfv	4
LIV : tmLwslsvssfavggmvasffggwlg	11

Comparison of leucine-zipper motif regions from different glucose transporters. CEF, chick embryo fibroblast; Ftlsk, human glucose transporter-like protein from human fetal skeletal muscle (type 3); HepG, glucose transporter from HepG2 human hepatoma cells (type 1); RatB, rat brain glucose transporter (type 1); IRGT insulin-regulatable glucose transporter from rat adipocyte (type 4); LIV, glucose transporter from rat liver (type 2). The leucine (L)/isoleucine (I) residues of the proposed zippers are in upper case; double dots indicate the one case where I replaces L; single dots, conserved amino acids.

least four different types of glycoprotein with 12 membrane-spanning domains which differ in tissue distribution and are encoded by discrete genes. As shown in the figure, the leucine-zipper motif occurs in the glucose transporter from chick embryo fibroblasts, which we have partially sequenced, and in all the other published transporter sequences except those from the liver. The sequence similarity in this region is striking considering that the overall sequence similarity between different types of transporters is around 65-70 per cent.

The idea that the glucose transporter dimerizes is attractive for several reasons. First, much direct evidence indicates that glucose-transport activity is associated with a protein that corresponds to the size of a dimer³. Second, radiation-inactivation data indicate oligomerization of the functional glucose transporter *in vivo*⁴. Third, formation of a heterodimer between the insulin-regulatable glucose transporter of insulin-sensitive tissues⁵ and brain types of glucose transporter experimentally introduced into such cells could explain why the latter can be regulated by insulin^{6,7}: the heterologous transporter could then ride 'piggy-back' with

the endogenous one. Finally, heterodimerization could similarly explain how, when a human transporter⁸ is introduced into chick embryo fibroblasts, its turnover is regulated in the same way as the endogenous transporter (unpublished data).

MARTYN K. WHITE
MICHAEL J. WEBER

Department of Microbiology,
University of Virginia Health Sciences
Center,
Charlottesville,
Virginia 22908, USA

1. Landschultz, W.H., Johnson, P.F. & McKnight, S.L. *Science* **240**, 1759–1764 (1988).
2. Buckland, R. & Wild, F. *Nature* **338**, 547 (1989).
3. Langdon, R.G. & Holman, V.P. *Biochem. biophys. Acta* **A945** 23–42 (1988).
4. Cuppoletti, J., Jung, C.Y. & Green, F.A. *J. biol. Chem.* **266**, 1305–1306 (1981).
5. James, D.E., Strube, M. & Mueckler, M. *Nature* **338**, 83–87 (1989).
6. Gould, G.W. et al. *J. biol. Chem.* **264**, 2180–2184 (1989).
7. Asano, T. et al. *J. biol. Chem.* **264**, 3416–3420 (1989).
8. Mueckler, M. et al. *Science* **229**, 941–945 (1985).
9. Kayano, T. et al. *J. biol. Chem.* **263**, 15245–15248 (1988).
10. Birnbaum, M.J., Haspel, H.C. & Rosen, O.M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5784–5788 (1986).
11. Thorens, B., Sarkar, H.K., Kaback, H.R. & Lodish, H. *Cell* **55**, 281–290 (1988).
12. White, M.K. & Weber, M.J. (unpublished data).

Multiple sclerosis and paramyxovirus

SIR—Vandvik and Norrby¹ used a different technique to the ones we described². They show the presence of intrathecally synthesized antibodies to the simian paramyxovirus SV5 (or related viruses) in the cerebrospinal fluid (CSF) of some multiple sclerosis patients. Contrary to our observations, however, none of the CSF samples contains significant amounts of such antibodies.

We have carried out further studies (with the help of B. Souberbielle and others) on the CSF samples from an additional 23 patients. Of these, eight contained SV5 (or related) antibodies but none had significant amounts of such antibody in their oligoclonal immunoglobulin G bands as judged by the effect of virus adsorption. Further experiments have also suggested that the relatively cathodic oligoclonal immunoglobulin Gs are very sensitive to changes in phosphate ion concentration and in pH, and it may be that some of the adsorptions noted in the original series of experiments were due to non-immune reactions. Nevertheless, not all our results (Fig. 3, ref. 2, for example) can be explained in this way. We conclude that antibodies to SV5 or related viruses constitute a major component of the CSF in some multiple sclerosis patients but that this proportion is much smaller than was originally claimed².

In some of the CSF samples examined by Vandvik and Norrby the oligoclonal bands revealed by blotting onto SV5 antigen cross-react either with PF-2 or with mumps. Other studies³ have shown

that such cross-reactions are characteristic of only a small proportion of the total antigen repertoire of these viruses and are mostly confined to internal antigens. This observation would be consistent with a persistent paramyxovirus infection⁴. Alternatively, it could be indicative of the involvement of another uncharacterized paramyxovirus elaborating only cross-reacting antibodies, the other more specific ones not being detected.

W.C. RUSSELL
R.E. RANDALL

Department of Biochemistry and
Microbiology,
University of St Andrews,
St Andrews,
Fife KY16 9AL, UK

K.K.A. GOSWAMI

Virology Section,
Department of Medical Microbiology,
University College and Middlesex
Hospital Medical School,
London W1P 7LD, UK

1. Vandvik, B. & Norrby, E. *Nature* **338**, 769–771 (1989).
2. Goswami, K.K.A., Randall, R.E., Lange, L.S. & Russell, W.C. *Nature* **327**, 224–247 (1987).
3. Randall, R.E. & Young, D.F. *J. gen. Virol.* **69**, 2051–2060 (1988).
4. Randall, R.E. & Russell, W.C. in *The Paramyxoviruses* (ed. Kingsbury, D.) (Plenum, New York, in the press).

Validity of sexual selection in birds

SIR—Read and Harvey¹ suggest that a reanalysis of comparative data on the relationship between avian plumage, brightness and parasite intensities does not substantiate our² or Read's^{3,4} previous findings that brighter species are more infected with haematozoa. I believe that the general finding of higher parasite prevalence in brighter, more sexually selected species remains valid.

First, I question the use of a two-tailed statistical test of the correlation. Read and Harvey¹ state, "we believe that the null hypothesis would have been rejected had a significant negative relationship been found". In view of our clear statement of our hypothesis about interspecies differences, I cannot imagine how a significant negative correlation could be construed as supporting it. Similar recognition of the one-sided nature of the hypothesis appears both in Read's analysis of haematozoa in North American and European passerines³, and in his review of recent research in the field of parasites and sexual selection⁴. The point is a critical one, because the use of one-tailed tests renders nearly all of the across-species correlations using the scores of the six respondents used by Read and Harvey significant at the 0.05 level, and the mean correlation significant at $P < 0.05$ as well. (Given the extremely non-normal distribution of the scores, it might be argued that the mean is not the best

summary statistic to use, and a modal score might be more appropriate, along with a non-parametric test.) The reanalysis of Read and Harvey¹ thus reveals differences in degree rather than kind. When these results are considered together with the finding for European birds² ($P = 0.002$) the overall case for an association remains very strong.

Nonetheless, the reported correlations between parasite prevalence and brightness, whether still significant or not, are generally lower than those calculated by us² or by Read³. In the re-ratings there may have been less emphasis on structural and display characters than I applied. Boat-tailed grackles, for example, received lower scores than common grackles from several of the respondents and none rated them higher (A. Read and P. Harvey, personal communication). I did rate them higher by one point; the common grackle is perhaps more iridescent but lacks the exaggerated tail which gives the former species its name. In ranking the original species², I gave species 'credit' for characters such as long tails or crests. This, together with unavoidable interest in forms of showiness that might reveal health may explain some of the discrepancies in scoring. In addition, my scores relied only partly on field guides and were constructed from personal experience of seeing virtually all the species in the wild. This perspective may give a more accurate version of the 'true' brightness of the species than merely using paintings or photographs from field-guides, which, as many birdwatchers know, can be misleading concerning the appearance of a bird under natural conditions. The real point is not showiness but the ability of females to detect health⁵.

I am also puzzled by the rejection of Baker and Parker's method for evaluating showiness. Use of this technique by Read³ resulted in a highly significant positive correlation between brightness and parasite prevalence of 113 European passerine species. In Read and Harvey's reanalysis¹, however, the method of Baker and Parker is dismissed on the grounds that the originators of the method "had their own view concerning the evolution of bright colours". But the important point is not whether the individual scorer has any ideas about the evolution of bird coloration, because many people (including, perhaps, the six ornithologists used as scorers by Read and Harvey) have some notions on the matter, but whether the scoring was performed blind, with the scorer ignorant of any associations between the variable in question and the items being scored. All that one can strive for in the present case is that the scorers are ignorant of the parasite data, which was true both for Baker and Parker, as their paper preceded ours, and for our

original work. As scorer, I might have been influenced by Hamilton's enthusiastic mention of particular species in early stages, and he had seen some of the parasite data, but several of the data sets used were added to the analysis after I had ranked the birds, and thus the final associations were truly unknown.

A final point concerns the finding that rarely-trapped birds show the correlation between brightness and parasites more strongly than commonly caught ones. I find this discovery intriguing, and can devise several explanations for it, including ones which support our hypothesis. For example, suppose that many different evolutionary forces, including the one suggested by us² have contributed to the showiness of birds. Birds might be rarely netted either because they are genuinely rare, or because for some reason they can usually avoid capture, either as a general rule or in the particular habitat in which the net is set. If sick birds, with large numbers of parasites, find it more difficult to elude the mist nets, and these tend to belong to showier species because of the operation of the mechanism we have suggested, then the relationship between rarity and the strength of the correlation reported by Read and Harvey is expected. Obviously other explanations can be constructed but as it stands, the result neither supports nor refutes our hypothesis. Likewise, the association with phylogeny is interesting, but has other possible explanations besides that of an "artefact".

Comparative work is rife with pitfalls⁷, and the possibility of previously unconsidered ecological or other factors confounding the results is always present. The generality of correlation between parasite prevalence and showiness still appears robust; the next challenges are to test the hypothesis within species⁸ and to elucidate more carefully the meaning of sexually selected characters to animals in their natural environments.

MARLENE ZUK

Department of Biology,
University of New Mexico,
Albuquerque,
New Mexico 87131, USA

READ AND HARVEY REPLY—Zuk claims that the comparative data we used to test the parasite theory of sexual selection² produces significant correlations between male brightness and parasite prevalence in passerine birds. That claim is wrong because it is based on cross-species correlations. It is well established that the non-independence of species points means that probability values attached to cross-species correlations are invalid unless phylogenetic associations have been statistically controlled for^{3,7,9-11}. The non-significant correlations presented in our original paper using analyses which

control for phylogeny are not altered by considering the prediction as one-tailed¹.

Whether there is justification for using one-tailed statistics, as Zuk claims that there is, will be relevant to future comparative tests of Hamilton and Zuk's theory². The question is whether both positive and negative associations between showiness and parasite load might evolve as a consequence of female choice based on showiness for resistant males. The answer hinges on whether female choice of resistant males increases resistance in a host population.

Hamilton and Zuk² assumed that it did not and they therefore predicted a positive correlation. If however, parasite load is reduced as a consequence of female choice for resistant mates, showiness might become negatively correlated with parasite load. The strength of natural selection against the evolution of bright coloration can be expected to differ among species, thus leading to different solutions to the conflict between the opposing forces of natural selection for dull coloration and sexual selection for bright coloration. Females of bright species will, according to Hamilton and Zuk's theory, more accurately assess resistance among males. Thus resistant males will obtain a greater mating advantage in brighter species, which will consequently harbour a lower parasite load. We know of no data which bear on the issue of whether female choice increases resistance in the host population. In the meantime, therefore, it seems prudent to consider the comparative prediction as two-tailed.

Zuk questions our reasons for re-scoring the showiness of European birds^{3,6}. As we stated¹, our failure to replicate the patterns found using Hamilton and Zuk's² scores of

the North American birds led us to investigate further the European birds. We had these latter birds re-scored because we wanted replicable scores by independent assessors. The best showiness criteria for such an exercise were surely those given by Hamilton and Zuk in their original paper², rather than those used for other purposes⁶.

As it stands, Zuk's suggestion for the low sample size effect fails to explain why parasite prevalence and brightness covary only in species in which only a few birds have been sampled for haematozoa.

Zuk¹ points out that in addition to the criteria given in the original paper⁸, she emphasized structural and display characters when scoring bird showiness. We are left wondering which characters were used, and how they were weighted so that the exercise can be repeated by other scorers. Zuk provides no reason for us to alter our original conclusion¹. Our comparative analyses do not provide general support for Hamilton and Zuk's theory.

ANDREW F. READ
PAUL H. HARVEY

Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK

1. Read, A.F. & Harvey, P.H. *Nature* **339**, 618–620 (1989).
2. Hamilton, W.D. & Zuk, M. *Science* **218**, 384–387 (1982).
3. Read, A.F. *Nature* **327**, 68–70 (1987).
4. Read, A.F. *Trends Ecol. Evol.* **3**, 97–102 (1988).
5. Endler, J.A. & Lyles, A.M. *Trends Ecol. Evol.* (in the press).
6. Baker, R.R. & Parker, G.A. *Phil. Trans. R. Soc. B* **267**, 63–130 (1979).
7. Pagel, M.D. & Harvey, P.H. *Q. Rev. Biol.* **63**, 413–440 (1988).
8. Zuk, M., Thornhill, R., Ligon, J.D. & Johnson, K. *Am. Zool.* (in the press).
9. Harvey, P.H. & Mace, G. in *Current Problems in Sociobiology* (ed. King's College Sociobiology Group) 343–362 (Cambridge Univ. Press, 1982).
10. Ridley, M. *The Explanation of Organic Diversity* (Clarendon, Oxford, 1983).
11. Felsenstein, J. *Am. Nat.* **125**, 1–15 (1985).

Cold fusion ideas

SIR—In the excitement about the possibility of chemically induced nuclear fusion^{1,2}, there is a danger that observations may be considered indicative of fusion when more conventional explanations from cognate fields are overlooked³. We wish to draw attention to two such cases.

First, consider an experiment in which identical currents are passed through identical electrolytic cells containing electrolytes differing only in the substitution of heavy water (D₂O) for H₂O, and the temperature of each cell is monitored. It has long been established that the viscosity of D₂O is appreciably higher than that of H₂O (1.26×10^{-4} Pa s and 1.009×10^{-4} Pa s at 20 °C, respectively)⁴. Making the assumption that H₂O/D₂O systems obey Walden's rule⁵, the resistivity of a salt solution in D₂O will be 25 per cent higher than the corresponding solution in H₂O; for a 0.1 M lithium acetate electrolyte solution, running at a current

density of ~ 0.1 A cm⁻² (ref. 2) between electrodes 1 cm apart, this will correspond to an excess heat generation in the D₂O solution relative to the H₂O solution of ~ 1.4 W per cm² of electrode surface. The effect should be still more marked with a strong acid or strong alkali (for example, LiOD) as electrolyte, where the conduction mechanism involves H/D tunnelling. We have checked that heat differentials of this order are in fact observed; we find that such cells typically run with D₂O 2–3 °C hotter than H₂O, with an overall temperature rise of 10 °C above ambient.

As a second example, tritium atoms should be produced in the proposed fusion reactions^{1,3}, and scintillation measurements of tritium decay have been reported². There have been suggestions that tritium has been detected mass spectrometrically as the ion DT⁺ (with a mass/charge ratio of 5 a.m.u.). In such an experiment the gas that diffuses through the walls of a closed palladium tube, which forms the cathode of an electrolytic cell, is

observed using a mass spectrometer. We have duplicated this experiment, and find that there are signals both at mass 5 and at mass 6; both are readily detected using a quadrupole mass spectrometer when the D_2 pressure in the system exceeds 10^{-9} mbar. These ions can also be generated, however, by admitting D_2 to the system. Ions of mass 5 and 6 are isotopically substituted H_3^+ ions, D_2H^+ and D_3^+ ; H_2 is always present as a background gas. The H_3^+ ion has a history as old as mass spectrometry itself^{6,7} and is well recognized as the product of reactions inside the mass spectrometer⁸, although the analysis has been carried out at pressures higher than those used here. The intensities of the various H_3^+ ions have nonlinear dependencies on pressure, and because the pressure increases in an electrolysis experiment, the concomitant intensity changes could be misinterpreted as the result of a time-dependent process.

M. E. BRIDGE
D. R. LLOYD

Department of Chemistry,

J. M. D. COEY

Department of Physics,
Trinity College,
Dublin 2, Ireland

1. Jones, S.E. *et al.* *Nature* **338**, 737–740 (1989).
2. Fleischmann, M. *et al.* *J. electroanal. Chem.* **261**, 301–308 (1989).
3. Pauling, L. *Nature* **339**, 105 (1989).
4. Moller, T. *Inorganic Chemistry* 393 (Chapman and Hall, London, 1952).
5. Atkins, P.W. *Physical Chemistry* 3rd edn, 677 (Oxford University Press, 1966).
6. Dempster, A.J. *Phil. Mag.* **31**, 438 (1916).
7. Thompson, J.J. *Rays of Positive Electricity* 32 (Longman, London, 1933).
8. Stevenson, D.P. in *Mass Spectrometry* (ed. McDowell, C.A.) 589 (McGraw-Hill, New York, 1963).

Delayed birth

SIR—In North America, the gestation period of the nine-banded armadillo (*Dasypus novemcinctus*) is usually 8–9 months, which includes a 3–4-month period of delayed implantation of the blastocyst before embryo growth. We recently reported that 20 armadillos captured during the period when blastocysts usually implant (November–February) underwent embryonic diapause of an additional year, which were probably induced by stress of capture¹. In one case, however, the embryos remained dormant for two years.

We have now observed a prolonged diapause in a second animal. The female was captured in November 1986 and immediately isolated from males. She bore no young during 1987 or 1988. But in February 1989, 30–32 months after her last mating season in the wild, she produced a litter of three living males.

These observations leave no doubt that armadillo embryos can survive *in utero* for at least 2–3 years. Are multiyear delays in birth peculiar to armadillos, or can they occur in other mammals, particularly

those with a prolonged reproduction delay, such as the European badger (*Meles meles*), which normally has a diapause of 10 months, or the wolverine (*Gulo gulo*), which has a diapause of 6–8 months²? We urge anyone who studies newly caught wild mammals to look for this phenomenon.

ELEANOR E. STORRS

H. P. BURCHFIELD

Florida Institute of Technology,
Melbourne, Florida 32901, USA

1. Storrs, E.E., Burchfield, H.P. & Rees, R.J.W. *Leprosy Rev.* **59**, 11–15 (1988).
2. Mead, R.A. in *Embryonic Diapause in Mammals*, *J. reprod. Fert. Suppl.* **29**, 11–24 (1981).

Cell identity resolved

SIR—Stratton *et al.* recently pointed out¹ that the putative medulloblastoma cell line TE671 “is most probably a subline of the rhabdomyosarcoma cell line RD”. Having carried out a cytogenetic study, we support that view.

The TE671 cell line (ATCC HTB 139) came to the American Type Culture Collection (ATCC) in 1982² from the Naval Biosciences Laboratory, which received the cell stock from the originator's laboratory³ in 1977. ATCC's TE671 distribution stock was prepared and released as HTB 139 in 1983.

As a result of the report of Stratton *et al.*, we reinvestigated the cytogenetics of both HTB 139 and the seed and distribution stocks of the RD cell line (ATCC CCL 136)^{2,3}, the last of which was accessioned in 1969. The modal number of chromosomes in HTB 139 is 47 but 42 per cent of the cells have 48 chromosomes; the karyotypes are similar to those reported previously for the cell line^{3,4}. For both RD stocks, the modal chromosome count is 48. RD and TE671 have very similar karyotypes. A few markers are specific to each cell line but such differences are common in mixed cultures of different cell lines^{5,6}. In conclusion, TE671 is undoubtedly a derivative of the RD. A detailed cytogenetic study on both RD and TE671 cells, together with the characteristics of their DNA variable number of tandem repeats (VNTR), will be reported.

T. R. CHEN
C. DOROTINSKY
M. MACY
R. HAY

American Type Culture Collection,
12301 Parklawn Drive,
Rockville, Maryland 20852-1776, USA

1. Stratton, M.R. *et al.* *Nature* **337**, 311–312 (1989).
2. ATCC Catalogue of Cell Lines and Hybridomas 6th edn (1988).
3. McAllister, R. *et al.* *Int. J. Cancer* **20**, 206–212 (1977).
4. Friedman, H.S. *et al.* *J. Neuropath. exp. Neurol.* **42**, 485–503 (1983).
5. Chen, T.R. *et al.* *Cancer Genet. Cytogenet.* **27**, 125–134 (1987).
6. Chen, T.R. *Cytogenet. Cell Genet.* **48**, 19–24 (1988).

Origin of soil magnetite

SIR—Maher and Taylor¹ misinterpreted our previous work² when they concluded that the ultrafine-grained magnetite they extracted from soils could not have resulted from the activity of dissimilatory iron-reducing bacteria such as strain GS-15. Although the magnetite they extracted from soils resembles the magnetite produced during the metabolism of GS-15, they excluded dissimilatory iron reduction as the source for this magnetite on the basis that there were no bacterial cells associated with the soil magnetite. But the magnetite produced by GS-15 is extracellular and is not firmly attached to the cells. The technique that Maher and Taylor used to extract the magnetite from soil does not extract non-magnetic bacterial cells such as GS-15. Thus, the evidence that they present does not exclude the possibility that the production of the soil magnetite is the result of the activity of dissimilatory iron-reducing bacteria.

DEREK R. LOVLEY

Water Resources Division,
US Geological Survey,
Reston, Virginia 22092, USA

JOHN F. STOLZ

Department of Biochemistry,
University of Massachusetts,
Amherst,
Massachusetts 01003, USA

MAHER AND TAYLOR REPLY—As we stated, it is difficult to distinguish between soil-formed magnetites and biological (GS-15 type) magnetites, owing to their similar shape, purity and magnetic properties. Furthermore, as Lovley and Stolz observe, our work on the soil magnetites did not attempt specific biological analyses. Nevertheless, our identification of soil magnetite as a product of inorganic, soil-forming processes is substantiated by, first, our demonstration of a totally non-biological pathway of magnetite formation, which produces quantities of ultrafine-grained magnetite rapidly and easily, under conditions characteristic of the soil environment³; and second, by the occurrence of the soil magnetite in well drained, oxidized soil profiles, an environment inimical to the anaerobic biological activity reported for GS-15².

B. A. MAHER

School of Environmental Sciences,
University of East Anglia,
Norwich NR4 7TJ, UK

R. M. TAYLOR

CSIRO, Division of Soils,
Glen Osmond,
South Australia

1. Maher, B.A. & Taylor, R.M. *Nature* **336**, 368–370 (1988).
2. Lovley, D.R., Stolz, J.F., Nord, G.L. & Phillips, E.J.P. *Nature* **330**, 252–254 (1987).
3. Taylor, R.M., Maher, B.A. & Self, P.G. *Clay Miner.* **22**, 411–422 (1987).

The accidental talent

Efraim Racker

For the Love of Enzymes: The Odyssey of a Biochemist. By Arthur Kornberg. *Harvard University Press: 1989. Pp. 336. \$29.95, £23.95.*

A HUNDRED years ago, there were microbe hunters. Then came vitamin hunters, then enzyme hunters and now there are gene hunters, whose reign is being challenged by the neurobiologists. But Arthur Kornberg has remained an enzyme hunter and in this book he explains why: he has never met a dull enzyme and deplores the present tendency to ignore them. Why have our students lost interest in enzymes, detesting "purification"? He says that "without knowing and respecting enzymes, better still loving them, answers to the most basic questions of growth, development and disease will remain beyond reach". Is the cloning of genes really more intellectually stimulating?

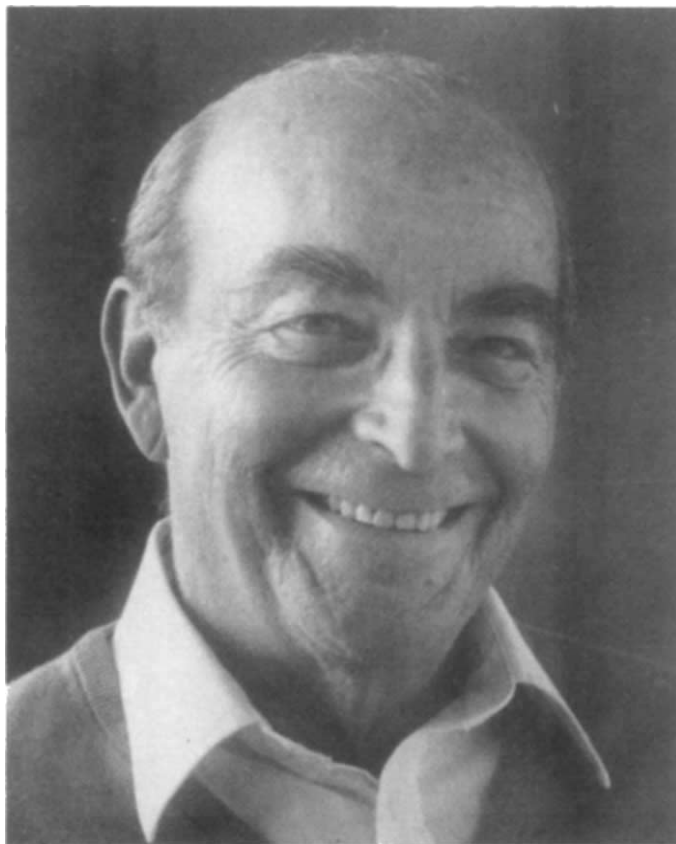
Kornberg's story of his research is interspersed with accounts of the history of medicine, personal experiences and his opinions about scientific life. It has wisdom, a sense of humour and a deep conviction of the importance of science in society. I share Kornberg's despair about the lack of understanding by the public and members of the US Congress of the need for basic science and the perhaps congenital inability of scientists to lobby for their cause.

Born into a poor immigrant family, Kornberg worked his way through New York City College, having a job as a salesman in the evenings and at weekends. In 1937 he was admitted to the Rochester medical school where, in spite of a brilliant record, he never received support from Dean G.H. Whipple, a Nobel laureate. Kornberg is still bitter about the "virus of anti-semitism" that pervaded Rochester and other medical schools.

Kornberg's first paper in 1942 — on jaundice — attracted the attention of R. Dyer, then director of the National Institutes of Health (NIH), who sent a team of medical officers to seek advice from Whipple on the jaundice observed in recruits vaccinated against yellow fever. Imagine Whipple's astonishment, Kornberg recalls with obvious delight, when the officers also asked to see Arthur Kornberg, a medical intern, the other 'authority' on jaundice at Rochester.

Shortly thereafter, Kornberg joined the US Navy as a ship's doctor, but was transferred in the autumn of 1942 to the Nutritional Laboratory of NIH. By 1945 he was bored with feeding rats with purified diets and went to work for 18 months in the laboratories of Ochoa and Cori.

I have often recounted an incident that took place during Kornberg's stay with



Arthur Kornberg — "in my theater . . . DNA and RNA provide the script, but the enzymes do the acting".

Ochoa, after both had worked for several weeks on the purification of malic enzyme from hundreds of pigeon livers. Late one evening, they finally came to the last step, precipitation of the enzyme with ethanol. They collected the fraction supposed to contain the enzyme in a cylinder, when Kornberg inadvertently spilled it on the floor. Kornberg was very, very upset but, the next morning, he noticed that the supernatant, kept at -15°C , had become turbid. He centrifuged it, collected the precipitate and tested it. It contained the bulk of the activity, being several times more active than the best of previous preparations. The 'new procedure' was duly published (without reference to the spilled cylinder).

What are the lessons of this story? First, a good enzymologist never discards any fraction until he recovers at least 120 per cent of the original activity (given that there are inhibitors in crude extracts). Second, Kornberg is clearly a prince of serendipity, who according to the legend has the gift of making important discoveries by accident. Researchers such as Edward Buchner (the discoverer of cell-free fermentation) who make one accidental discovery in their lives are the ordinary citizens of science; those who are accident prone, like Kornberg, are the princes.

Kornberg's account of his many adventures with enzymes is unfailingly lucid. He tells how he searched for the "best potato" to isolate the enzyme that cleaves NAD,

of his excitement on discovering that yeast is a good source for enzymes involved in the synthesis of NAD, and how luck struck again when he discovered an enzyme for NADP formation: he used a crude ammonium sulphate fraction, not a crude extract containing inhibitors. This was a lesson Kornberg remembered when he embarked on the trail of DNA. In studying the synthesis of the precursor orotic acid, he had the "lucky thought" of mixing extracts from mammalian (liver) and yeast cells (one of my favourite reconstitution cocktails) and found a hundred-fold amplification in activity, obviously because each extract had a different rate-limiting enzyme in the pathway. The liver enzyme led him on to the discovery of 5-phosphoryl-1-pyrophosphate (PRPP) and of PRPP synthase, and eventually to the synthetic pathway of UTP, CTP and later of ATP and GTP.

These historic discoveries encouraged Kornberg to pursue the synthesis of DNA and led to the discovery of

DNA polymerase. Recalling his various love affairs with enzymes at his 70th birthday celebrations, Kornberg described that with DNA polymerase as the longest and most passionate because he did so many of the experiments with his own hands. The message of this is clear: many young scientists now leave the bench too early, robbing themselves of the special joy of discovery.

DNA polymerase was not easily conquered. With only a few radioactive counts a minute above background, purification was hard work. But finally the proof was in: in the presence of DNA template and of all four-deoxyribose derivatives of ATP, GTP, CTP and TPP, DNA was synthesized as a faithful copy of the added

primer. The two papers describing this discovery were first rejected by the *Journal of Biological Chemistry*, but rescued by the then new chief editor, John Edsall. The Nobel prize awarded to Kornberg (with Ochoa) in 1959 did not turn his head; he returned from Stockholm with his enthusiasm for research undiminished.

There was great excitement when DNA polymerase duplicated an infectious $\phi\chi$ DNA, but the story was blown out of proportion by the press. When it later became clear that this is not how DNA is synthesized, but how it is repaired, some complained that Kornberg had overstated his case. Although probably painful at the time, it stimulated Kornberg to start from the beginning in his search for the mechanism of DNA replication.

Kornberg picked a new system, the filamentous phage M_{13} (a lucky choice, he says), and persuaded his entire group to join him in the adventure. They found that RNA synthesis is required for DNA duplication, but from then the story became more and more complex. After 17 years of work on DNA synthesis, Kornberg is still "grappling with its awesome complexity".

Over a dozen proteins were discovered by methods of biochemical resolution and genetic complementations. He predicts that over 20 components are required and that it will take decades before we fully understand the DNA-replicating machinery. Year after year new discoveries are being made and even certain phospholipids, capable of releasing tightly bound and inhibitory ADP from DNA A protein, appear to be involved. Could such a mechanism have a part in some of the G proteins that tightly bind inhibitory GDP?

Kornberg is an innovator in more than his own research. His department at Stanford, where all share funds, fun and discoveries, is unique. He was, and is, a demanding mentor. At NIH, Bill Price, his gentle technician, once failed properly to balance a set of centrifuge tubes and the tube broke. Kornberg chided him that "the hour lost can never be recovered". On another occasion, when Kornberg had failed to understand a Brazilian post-doctoral fellow, he asked him what he was driving at. The response was, "Me driving? No! You drive! You drive all the time!"

Kornberg's opinions of the conduct of science are scattered through his book. He believes, for example, that the role of NIH in the history of medical science has been undervalued, saying that "the achievements of NIH defy exaggeration". Members of Congress would do well to read the passage concerned (pp.129–135), and do more to preserve that endangered species, the NIH researcher.

On the atmosphere of secrecy emerging from the inevitable and potentially beneficial marriage between industry and the

universities, Kornberg takes a paradoxical line: "I believe that secrecy makes even less sense in an industrial setting than in an academic one. The best insurance for industrial success is an open atmosphere that provides optimal access to all information and advice. As significant advances are made, inventions can be protected each step along the way by a patent-applications team that interacts on the spot continually and knowledgeably with the scientists. In contrast, a company policy of secrecy closes the avenues of exchange with academic scientists and prevents useful collaborations with other companies. It is generally unappreciated that secrecy within the company itself shelters mediocrity and discourages the vigorous exchange that identifies excellence". Will his cogent arguments persuade the leaders in industry? I wish I could be optimistic about this. But they might heed his other suggestion, that a partnership between academic science and industry might be effective in lobbying Washington

for support for basic science.

Kornberg is a gifted writer and the book contains sentences of both wisdom and beauty. Thus (on p.274): "The tides of fashion in science erode one beach to create another. In the rush and excitement of the new mastery over DNA, training in enzymology and its practice have been neglected. Most of our students are introduced to enzymes as commercial reagents, and they find enzymes as faceless as buffers and salts". Or (p.298): "... DNA and genes captured the spotlight from enzymes; but in my theater, enzymes kept the leading role. DNA and RNA provide the script, but the enzymes do the acting".

For the Love of Enzymes is at the same time charming, stimulating and informative, a reflection of its author. If widely read — and I very much hope it is — it will greatly contribute to the scientific education of the public. □

Efraim Racker is Albert Einstein Professor of Biochemistry, Cornell University, Ithaca, New York 14853, USA.

On form

S.J. Perkins and S.A. Baldwin

Protein Structure: A Practical Approach.

Edited by T.E. Creighton. IRL Press: 1989. Pp. 355. Hbk £30, \$65; pbk £20, \$40.

THE *Practical Approach* series now has over 40 titles under its banner. This new addition, edited by T.E. Creighton, describes experimental techniques for the study of protein structures. Such structures are usually investigated because of the light they might throw on function, a property of proteins which is not discussed here (a companion volume is concerned with that topic). Moreover the book is largely concerned with globular, water-soluble proteins; both fibrous proteins (about half the protein in the human body!) and membrane proteins receive little mention.

Despite its title, the book also does not deal with the determination of protein primary structures (again, a topic covered by a different volume in the series). Nor does it deal with atomic structure determinations requiring sophisticated equipment. Instead, it presents a collection of simple techniques for the analysis of protein structure that could be performed in most biochemical laboratories. The description of each technique is prefaced by a brief introduction, implying the need for some background knowledge on the part of the reader. But the practical details are clear and will be easy to follow even for those with little experience of similar techniques.

Six of the 14 chapters are concerned

with various aspects of gel electrophoresis, including M_r estimations, isoelectric focusing, two-dimensional gel electrophoresis and peptide mapping. Scheidtmann gives a comprehensive account of the generation of antibodies to synthetic peptides for protein identification and characterization, while Friguet and colleagues provide a very helpful introduction to the immunochemical analysis of protein conformation.

Most of the remaining chapters also deal with protein conformation. In the light of the DNA-sequencing revolution, one of the most timely is that from Argos on structural predictions from sequences; this is a straightforward survey of how to determine a DNA reading frame, how to search for and align homologous sequences, and how to predict secondary structure and then plan site-specific mutagenesis experiments. There are also several contributions on the analysis of protein folding and stability, and on spectroscopic techniques for the determination of conformation and conformational changes. The coverage is fairly comprehensive, although omission of the increasingly popular technique of Fourier transform infrared spectroscopy for the determination of protein secondary structure is surprising.

Volumes in the *Practical Approach* series have become standard laboratory equipment. This one is no exception, and we must hope that the series stays on course under the new management of Oxford University Press. □

S.J. Perkins and S.A. Baldwin are in the Departments of Biochemistry and Chemistry, and of Protein and Molecular Biology, at the Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, UK.

Progress on the margin

R.A. Scrutton

Early Tertiary Volcanism and the Opening of the NE Atlantic. Edited by A.C. Morton and L.M. Parson. *Blackwell Scientific*: 1988. Pp. 477. £80, \$85.

THE Early Tertiary Igneous Province has been studied by geologists and geophysicists for decades, and is the subject of a voluminous, ever-growing literature. Long-standing research such as this usually proceeds at a measured pace, but occasionally receives a major fillip as a result of an unexpected finding.

In the late 1970s, discovery of 'dipping reflector sequences' on seismic reflection profiles from the continental margins of the North Atlantic, and their interpretation as thick volcanic sequences of early Tertiary age, indicated that the igneous province could be studied not only on the continents bordering the Atlantic, but also at the continent-ocean boundary whose formation was heralded by the volcanism. Marine geophysics and deep-sea drilling became important new tools. Further impetus to research came from the repeated discovery of volcanic rocks of relevant age in hydrocarbon exploration wells in basins on the north-west European continental shelf. The new information, together with re-interpretation of older data, provided a great advance in our understanding of volcanic processes at passive continental margins, and of their products on the margins and elsewhere.

The book reviewed here stems from a conference held in London two years ago (discussed in a News and Views article in *Nature* 327, 191; 1987), the object of the organizers being to highlight those processes involved in the inception of sea-floor spreading in the North Atlantic. The resulting volume contains contributions from geophysicists, geochemists, volcanologists, structural geologists and sedimentologists, among others, but perhaps more importantly it brings together work from onshore and offshore.

One of the outstanding problems of the volcanic province is to get the chronology of events right throughout the province — here the magnetostratigraphy and radiometric dating of continuously sampled volcanic sequences on land can be wedded with the biostratigraphy of the drilled but less continuously sampled volcanic-sedimentary sequences in basins offshore. The need for further work in this field is repeatedly mentioned by the contributors. A number of the outstanding questions on volcanic processes at passive margins are also aired. Are dipping reflectors only in oceanic crust or do they also

occur over continental crust? Is the asthenosphere upwelling process a passive one or does it involve local convection beneath the rift? Is the underlying mantle convection system an upper-mantle or whole-mantle phenomenon? An exciting feature discussed here is the discovery of peraluminous volcanic rocks in Rockall Trough and on the Voring Plateau, and their significance for the role of crustal and sediment melts in the volcanic province.

The Geological Society of London Special Publication series has become an important vehicle for conference proceedings. This, the thirty-ninth volume, preserves that reputation. Although not a large book, the coverage of the Tertiary province is broad, with petrology and

geochemistry providing about half of the contributions. With few exceptions the articles are of high quality, and some are excellent.

As a report on the progress of research in what has become the classical area for the study of volcanic passive margins, this book will be widely used and much referenced. For those interested in passive margins, in volcanic products and processes in extensional terranes, or in the general geology of the North Atlantic, it is a must. For others it provides a useful reference work at a wide range of levels. □

R.A. Scrutton is Reader in Marine Geophysics at the Grant Institute of Geology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK.

Hole process

Richard A.F. Grieve

Impact Cratering: A Geologic Process. By H.J. Melosh. *Oxford University Press*: 1989. Pp. 245. £45, \$65.

THE study of impact cratering is a relatively new enterprise for geologists. Much present interest stems from space exploration programmes, which firmly established cratering as an important process in early planetary evolution. More recently, the suggestion that the Cretaceous-Tertiary mass extinction was the result of a large impact has stimulated interest in — and debate about — impact as a process affecting the Earth.

Melosh's is the first English-language textbook on the subject. As such, it is a very welcome addition to the literature, bringing the subject together and filling in the gaps left by compendia of papers in conference proceedings volumes which have dealt only with selected aspects.

Research on cratering is a multidisciplinary pursuit which Melosh approaches from his own discipline of physics. This is the underlying theme and strength of the book. The author is frank on what is known and not known in this rapidly expanding field, explicitly stating, sometimes repeatedly, the assumptions and the caveats.

The reader is introduced to the subject in a logical fashion. I particularly liked the discussion of scaling relations coming before discussion of the processes by which craters are modified. It is easy to forget that experimentally derived scaling relationships generally refer to the excavation and so-called transient cavity, not to the final crater, which will have undergone considerable structural modification, particularly in large, complex craters.



From Impact Cratering

Dark side of the moon: the largest crater (centre) is about 75 km across.

The subtitle is rather misleading. This book is about processes, not the geology of impact craters. In fact, the geological effects of impact are given little space. For example, shock metamorphism is discussed in one page, whereas five are devoted to the process of viscous relaxation of crater morphology. There are also a few minor factual errors, particularly when the text deals with geological observations. A few descriptive words are given for each reference, at the ends of chapters, to help students wishing to go further.

Impact cratering is increasingly being recognized as an important geological process. This well-written and amply illustrated book will serve as an excellent reference and teaching text for senior undergraduate and graduate students. I certainly wish it had been available a few years ago, when I was trying to teach the subject. □

Richard A.F. Grieve is in the Geophysics Division of the Geological Survey of Canada, Building 3, Observatory Crescent, Ottawa, Canada K1A 0Y3.

In the film business

Andrew Holmes-Siedle

The Si-SiO₂ System. Edited by Pieter Balk. Elsevier: 1988. Pp. 356. Dfl. 220, \$115.75.

The Physics and Chemistry of SiO₂ and the Si-SiO₂ Interface. Edited by C. Robert Helms and Bruce E. Deal. Plenum: 1988. Pp. 556. \$138, £86.25.

ONE reason for the low cost of mass-produced microelectronics is the abundance of silicon on the Earth. Another reason, not so widely recognized, is the versatile nature of the oxide of silicon, silica (SiO₂).

When silicon is heated in oxygen or steam, a strong, tenacious form of silica grows as continuous film. This oxide film is a very useful working material for fabricating microelectronic structures. First and most important, the films are excellent insulators and are used as the oxide in the metal-oxide-semiconductor sandwich which is used widely in microelectronics. Second, the films can be patterned by photolithographic etching; this is what gives us the ability to make very fine patterns of p-n junctions by diffusion.

In both applications, the classical view was that silica was an inert spacer or barrier, unaltered by the application of fields or other stresses. The reality is different, and not surprising once we realize what a sensitive environment the silica has around it. Even tiny electronic changes in the silica are picked up in the sensitive layer of silicon underneath it and are reflected in the performance of the electronic device concerned. If the film fails, the device dies. Hence the reason why a large number of researchers are trying to understand the growth of silica films, their structure, and the electronic states which occur within the films and (especially important) at the interface between the film and the silicon wafer beneath.

Pieter Balk is a European authority on the silica film and in his book he has brought together a coherent set of contributions on the subject. By contrast, the volume edited by Helms and Deal is a snapshot of research in 1988. It contains over 60 short papers on SiO₂ and its interfaces that were delivered at a conference held in Atlanta in May 1988. In the classical view, insulators such as silica were not meant to carry currents at all unless total breakdown occurred. Contemporary research is defining how, in fact, charge enters, passes through and is trapped in the oxide film. An especially active region is the Si-SiO₂ interface, where charge-

trapping states have special ability to exchange electrons rapidly with the silicon. The standard of science required to unravel this picture is extremely high, despite the 'applied' nature of the problem.

The contributors to Helms and Deal present useful vignettes of the latest research but, despite the inclusion of some review papers, there is not much in the way of a coherent view of the subject. Some of the research papers are outstanding, however — for example those by Irene on the growth of oxides on silicon at high temperatures, by Bevk on the flatness or otherwise of the interface between single crystal Si and grown SiO₂ on the atomic scale, and by the late Frank Feigl and his school on charges in SiO₂. Damage to the silica film due to charge injection or irradiation is a very important subject and papers on this topic occupy a whole section of the book. Different but equally fascinating approaches to damage effects are taken by Edwards, Schulz and Ma. No less important are a contribution by Deal, which attempts to define what constitutes 'quality' in this use of silica, and one by Di Maria on charge transport in silica.

There can be few better people to organize a summary of insulator films than Balk. He himself contributes a long and masterly introductory section to *The*

Si-SiO₂ System. Four chapters are written by European authors, who are not only of high scientific standing but also explain themselves very clearly. Another is by Robert Helms, an editor of the other book reviewed here. Although many of the authors have spent long periods working in the United States, Balk's book demonstrates Europe's strength in the thin-film dielectrics field — a fortunate state of affairs considering the need for this skill in a competitive industry. The practical aspects of the use of silica films in microelectronics are nicely integrated with the explanations of the complex physics of the system.

In an apparently off-hand comment, Balk says that "The prescription for optimising MOS properties is basically a very straightforward one. The SiO₂ should be as perfect a network as possible...". This remark in fact encompasses neatly the whole complex and vital activity which forms the topic of these books. How well this prescription is met may make the difference between success and failure in one of the most crucial industries of the twentieth and (perhaps) the twenty-first centuries. □

Andrew Holmes-Siedle is owner of REM Instruments, 64A Acre End Street, Eynsham, Oxford OX8 1PD, UK, and a consultant to the Fulmer Research Laboratories, Stoke Poges, near Slough.

Off target

Tony Davies

Antibody-Mediated Delivery Systems. Edited by John D. Rodwell. Marcel Dekker: 1988. Pp. 384. \$99.75.

HYBRIDOMAS growing *in vitro* or *in vivo* can generate large quantities of antibody molecules with a wide range of binding specificities. That capacity has acted as a catalyst to the concept that drugs can be taken to their targets on the back of carrier molecules. The broad strategies are now well known in that an antibody molecule, that can specifically bind to a target such as a tumour, is covalently linked to another and cytotoxic agent. Such heterobifunctional conjugates, it is argued, will kill their targets without serious non-specific side effects.

The present volume is a collection of papers describing these possibilities. There is no editorial summary, nor apparently has any serious attempt been made to avoid repetition at the beginning of each chapter. On the positive side the accounts of materials and methods are relegated to a technical appendix in each paper, thus preserving the notion that review is intended. In fact the contributions are in most instances fairly standard

and presumably unrefereed research papers, which mainly portray the interests of their authors rather than attempting to appraise a particular aspect of the field.

Various conjugates of conventional chemotherapeutic agents with antibodies are considered, but although the experimental results are promising — as they always have been — the extension into clinical practice seems to be slow to mature. Some of the more esoteric approaches using, for example, cobra venom factor are portrayed — at least those tried *in vitro*. Frankel presents a useful review of linkage of proteinaceous toxins to antibodies. There is even a paper (for no obviously good reason) on methods of labelling with ⁶⁷Cu using porphyrins.

By picking and choosing, it is possible to find in this volume the states of various branches of the targeting art, but it is rather hard work. If, as is threatened, this is to be the first of a series of books on roughly the same topic, much more care should be given to selection and editing of papers so that the standards of critical review and redaction that readers have every right to expect are fulfilled. □

Tony Davies is Deputy Director of the Institute of Cancer Research, Haddow Laboratories, 15 Cotswold Road, Belmont, Sutton, Surrey SM2 5NG, UK.

Hipparcos: astrometry from space

M. A. C. Perryman

The Hipparcos satellite, due to be launched aboard Ariane on 27 July, aims to provide an enormous advance in the quality and quantity of stellar astrometric data.

HIPPARCOS is the European Space Agency's astrometry mission, a space experiment due to be launched on 27 July and dedicated to the precise measurement of the positions, parallaxes and proper motions of the stars. Over a planned lifetime of 2.5 years, the goal is to measure the astrometric parameters of about 120,000 stars with a precision of some 2–4 milliarcseconds, and the astrometric and two-colour photometric properties of 400,000 further stars with somewhat lesser precision (see Table 1).

More than twenty years of planning have gone into Hipparcos¹. A feasibility study by the European Space Agency (ESA) in the late 1970s led, in 1980, to a decision that the mission should be part of the agency's mandatory scientific programme, to which all member states must contribute. Development of the hardware began early in 1984, the satellite was completed and placed in storage in May 1988 and is now awaiting launch by an Ariane 4 rocket from French Guiana.

What makes the Hipparcos mission unique, and the expected results potentially so dramatic, is that the satellite can observe the entire celestial sphere from beyond the perturbing atmosphere, that the measuring instruments will be free from instrumental flexure under gravity, and that the thermal environment will be actively controlled. Differential angular measurements of the stars in the catalogue will be made over large angles, at many different orientations and at many different epochs. As a result, the parallaxes will be absolute, while regional or systematic errors in positions and annual proper motions are expected to be much less than a milliarcsecond.

The eventual outcome will be a rotation-free, or quasi-inertial, representation of the positions of the stars, constructed by relating the final Hipparcos catalogue to observations of extragalactic objects². The representation of the catalogue stars will also be linked to the radio reference frame by direct or indirect observations of radio stars or quasars by Very Long Baseline Interferometry (VLBI)³.

Objectives and outcomes

Optical astrometry has two main objectives. The first is to provide a non-rotating stellar reference frame to which the motions of objects in the Solar System and stars in the Galaxy may be referred, and by which optical observations can be related to those in other regions of the electromagnetic spectrum. Second, astrometry seeks to provide, within such a framework, basic observational data for the study of stellar properties (luminosity, mass and the like), the spatial distribution of stars within the Galaxy and their motion.

The position of an object in space may be described by its direction cosines with respect to some coordinate system at some epoch, its heliocentric distance and the rates of change of these quantities with time. Relative angular measurements made from astronomical observatories at well-separated epochs yield, in principle, both the direction cosines and their time dependence. If the distance of the star is known, the angular 'proper motion' can be converted into a velocity perpendicular to the line of sight.

A sequence of observations of an object, relative to more distant background stars, as the Earth travels in its annual orbit around the Sun, also yields the distance to the object by the determination of its parallax. Thus angular measurements define five astrometric parameters of the object: two positional coordinates, their rates of change with time (or proper motions) and the parallax (which is inversely proportional to the

distance). For a complete description of the motion of the object in space, assumed to be rectilinear and uniform, the motion along the line of sight, or radial velocity, must also be determined, most conveniently by Doppler measurements.

Although these quantities are fundamental, the angles involved are extremely small (parallaxes and annual proper motions are typically measured in milliarcseconds), so that the measurements required to quantify them must be extremely precise. Conditions at ground-based observatories — the atmosphere, the lack of all-sky visibility and the gravitational and thermal flexure of telescopes — complicate the measurements and confuse the separation of the positional and time-dependent parameters. The result, in recent years, is a tendency to suppose that there is a limit beyond which further progress towards greater astrometric precision will be technologically complex and observationally time-consuming. An individual positional measurement determined by automatic meridian circles, for example, has a mean square error of 0.18–0.20 arcsec.

Just as the scientific objectives of optical astrometry may conveniently be separated into the construction of the reference frame and the physical interpretation of parallaxes and proper motions, so two broad categories of ground-based measurements have evolved. Thus the present strategy for the elaboration of the reference frame rests on large-angle techniques; a fundamental frame based on meridian observations of a few thousand bright ($m < 8$ magnitude) stars is supplemented by meridian observations of further reference stars, reduced differentially with respect to the fundamental stars, and further interpolated by photographic surveys. Photographic parallax determinations, by contrast, are essentially small-angle measurements.

The resulting catalogues are of two kinds. There are small, but precise, catalogues such as the FK5 Catalogue of a few thousand stars with mean positional error at its central epoch (1949) of 0.019 arcsec, and there are large but less precise catalogues. An extreme example of the latter is the nearly complete Guide Star Catalogue developed for use with the Hubble Space Telescope (HST), comprising some 20,000,000 objects with external coordinate errors of 1–2 arcsec.

A complete separation of goals or of measurement techniques is neither feasible nor desirable. The accuracy of the reference frame deteriorates with time, because of errors in the determi-

TABLE 1 Summary of expected results

Main experiment:	Number of stars	120,000
	Limiting magnitude	$V = 12.4$ mag
	Complete to	$V = 7.3-9.0$ mag*
	Positional accuracy	0.002 arcsec ($B = 9$ mag)
	Parallax accuracy	0.002 arcsec ($B = 9$ mag)
	Proper motion accuracy	0.002 arcsec per year ($B = 9$ mag)
	Systematic errors	< 0.001 arcsec
Tycho experiment:	Number of stars	$> 400,000$
	Limiting magnitude	$B = 10-11$ mag
	Positional accuracy	0.03 arcsec ($B = 10$ mag)
	Photometric accuracy	0.05 mag in B and V (per observation)
	Observations per star	~ 100

* Depending on galactic latitude and spectral type.

nation of individual space motions (see Fig. 1), whereas small-scale angular measurements typically give relative rather than absolute parallaxes. As a consequence, the errors of parallaxes or of proper motions are affected by the external errors of the reference system to which the measurements are related.

Inevitably, recent developments promise substantial improvements of precision¹. Thus, the projected ground-based interferometric measurement of large angles and the astrometric use of the HST will also yield positional accuracy of a few milliarcseconds. But these measurements will usually be confined to a few celestial objects and their immediate locality. By contrast, the Hipparcos measurements will span the whole sky.

Measurement and design

The payload (Table 2; refs 5–9) is centred around an optical all-reflective Schmidt telescope (Fig. 2), with a novel beam-combining mirror that brings to a common focal surface the light from two fields of view, separated by about 58° and each of dimension $0.9^\circ \times 0.9^\circ$. As a consequence, it is possible to make both large-field and small-field measurements simultaneously.

The satellite sweeps out great circles over the celestial sphere, and the star images from the two fields of view are modulated at the focal surface by a highly regular grid of 2,688 transparent parallel slits covering an area of $2.5 \times 2.5 \text{ cm}^2$. The fabrication of the aspheric asymmetric beam-combining mirror and of the highly accurate modulating grid (by electron-beam lithography) were two of the key technical challenges in the development of the payload.

The satellite will spin slowly, completing a full revolution in just over two hours, but there will also be a continuous slow change of direction of the axis of rotation. Thus the telescope will scan the complete celestial sphere several times during its lifetime. In the process, starlight is modulated by the slit system and the modulated light is sampled by an image-dissector tube detector at a frequency of 1,200 Hz. At any time, there will be four or five of the selected (or programme) stars in the combined fields of view. The sensitive area of the detector is only small (covering an area of about 38 arcsec in diameter projected on the sky), so that it can follow the path of only one star at a time, but rapid computer control will allow it to be switched briefly to all the programme stars during their passage across the field, which takes about 20 seconds.

The telescope will continually determine the relative (along-scan) positions of the programme stars, each of which will

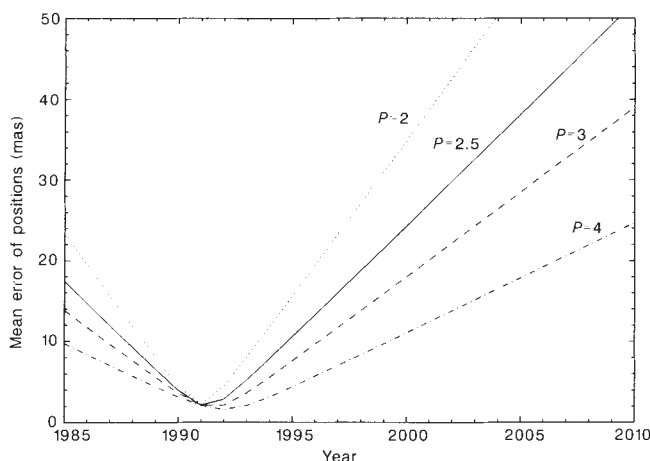


FIG. 1 The accuracy of the Hipparcos positional reference frame as function of time and for different mission lengths ($P = 2, 2.5, 3$ and 4 years). A mean error of 2 milliarcseconds (mas) is assumed for the positions at the mean epoch of the nominal mission ($P = 2.5$ years). Notice that the accuracy five or more years beyond the beginning of the mission is entirely determined by the proper motion standard error (equal to the slope of the almost straight lines), which would improve as $P^{-3/2}$.

appear first in the preceding and then in the following field of view because of the rotation of the satellite, so that several comparisons with different stars can be made. Because the axis of rotation of the satellite changes on each sweep of the sky, the same stars will reappear, but now compared with other stars. Thus a dense net of measurements of the relative angular separations of the stars is progressively built up.

The payload includes two star mappers (one redundant) to enable the precise attitude of satellite to be determined in real time (on-board the satellite) as well as *a posteriori*, on the ground. The same data are also used by the Tycho experiment to give astrometric and two-colour photometric data for about 400,000 stars down to about 10–11 magnitude.

Each star mapper consists of a non-periodic modulating grid (mounted alongside the primary grid) comprising two sets of four slits, each at a different inclination to the scanning direction. Two photomultipliers will measure the light transmitted by the whole star-mapper grid in two different spectral bands, roughly corresponding to the Johnson *B* (blue) and *V* (visual) bands. The differently inclined slits will allow the satellite attitude to be derived from the detector signals as the star images move across the grid. The digitized version of the modulated light signal from the photomultipliers will be sampled at 600 Hz.

Telemetered photon counts from the main detector and attitude information from the star mappers, together with other house-keeping data, will allow the relative phases of the star images in the combined field of view to be determined. The data processing will be carried out on the ground and will yield the final catalogue of star positions, parallaxes and proper motions.

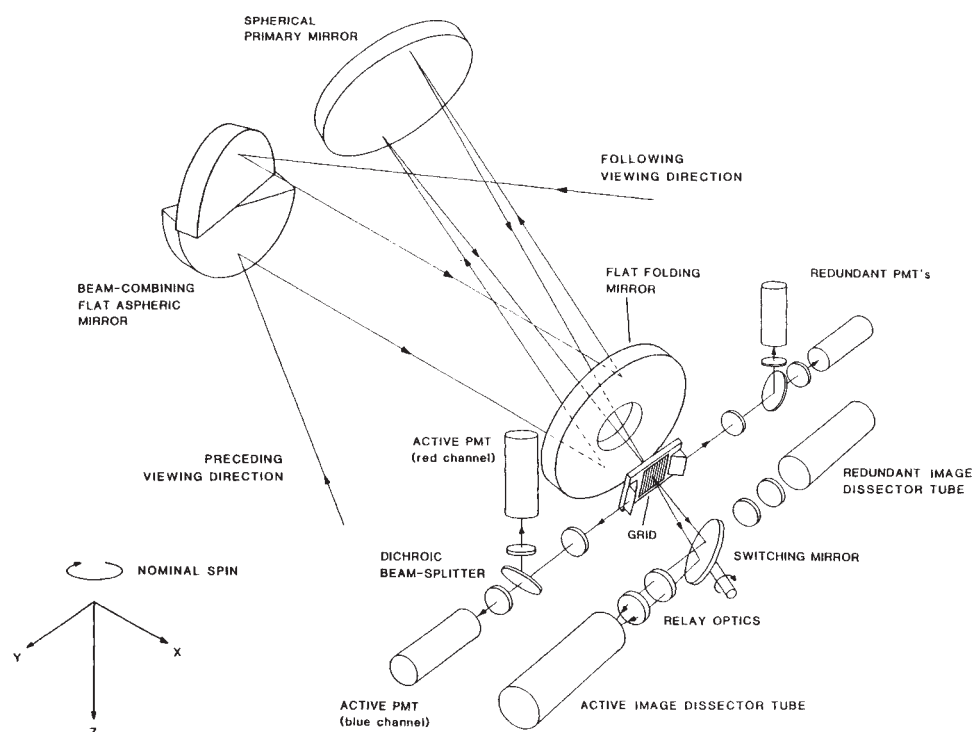
Data reduction

The satellite will produce some 24 kbits of data per second, so that about 10^{12} bits of data will be generated over the lifetime of the satellite, corresponding to some 150×10^6 'grid coordinates'. The problem of data reduction is essentially that of solving a

TABLE 2 Payload and satellite characteristics

Optics:	Telescope configuration	All-reflective Schmidt
	Field of view	$0.9^\circ \times 0.9^\circ$
	Focal length	1,400 mm
	Separation between fields	58°
	Diameter of primary mirror	290 mm
Primary detection system:	Mirror surface accuracy	$\lambda/60$ rms
	Modulating grid	2,688 slits
	Slit period	1.2 arcsec (8.2 μm)
	Detector	Image dissector tube
	Photocathode	S20
Star mapper (Tycho) system:	Sensitive field of view	38 arcsec diameter
	Spectral range	375–750 nm
	Sampling frequency	1,200 Hz
	Modulating grid	4 slits perpendicular to scan
	Detectors	4 slits at $\pm 45^\circ$ inclination
Satellite parameters:	Photomultiplier tubes	Bialkali
	Spectral range ('B')	$\lambda_{\text{eff}} = 430 \text{ nm}$, $\Delta\lambda = 90 \text{ nm}$
	Spectral range ('V')	$\lambda_{\text{eff}} = 530 \text{ nm}$, $\Delta\lambda = 100 \text{ nm}$
	Sampling frequency	600 Hz
	Launch mass	1,140 kg
	Power requirements	295 W
	Uplink data rate	2 kbit s^{-1}
	Downlink data rate	24 kbit s^{-1}
	Satellite orbit	Geostationary
	Spin axis inclination to Sun	43°
	Spin rate	168.75 arcsec s^{-1}

FIG. 2 The layout of the Hipparcos payload.



very large, sparse and over-determined system of equations with a large number of unknowns. In addition, some 10^6 attitude unknowns and some 20,000 instrumental unknowns (essentially the time-, position- and colour-dependent coefficients representing the geometrical properties of the measurement system) must either be eliminated or solved for.

The problem is too substantial to be treated directly, but it lends itself to linear decomposition by the formation of certain intermediate quantities — the star abscissae along the great circles swept out by the scanning motion of the satellite. Whereas more rigorous methods for solving the problem of the attitude and sphere reconstruction are still being studied, the two data-reduction consortia responsible for the work have taken as their starting point decomposition into the following three separate steps¹⁰⁻¹⁴ (Fig. 3):

■ The image-dissector tube photon counts are fitted to a star intensity model of the general form

$$I = B + I_0[1 + M_1 \cos(\omega t + \psi_1) + M_2 \cos(2\omega t + \psi_2)]$$

where B is the background, I_0 the unmodulated star intensity, ω the angular frequency of light modulation by the grid, M_1 and M_2 the first and second harmonic instrumental modulation coefficients and ψ_1 and ψ_2 the unknown signal phases. The phases are treated in data segments corresponding to a 'reference great circle', of about 12 hours' duration, chosen with celestial pole close to the mean position of the instrument's spin axis over the period. (Some 2,000 reference great circles will be generated in the 2.5-year lifetime planned.) The star abscissae, the along-scan attitude parameters and a set of instrumental parameters describing the geometric transformation between the celestial sphere and the measurement system, are solved independently for each reference great circle by a least-squares process.

■ The resulting star abscissae for a 'well-behaved' subset of some 40,000 of the programme stars are combined in a subsequent least-squares process, referred to as the sphere solution, which integrates all the reference great circles into a single global system. In this step, the abscissae are taken as given, and the astrometric parameters of the same stars, together with a single abscissa zero point for each reference great circle, are

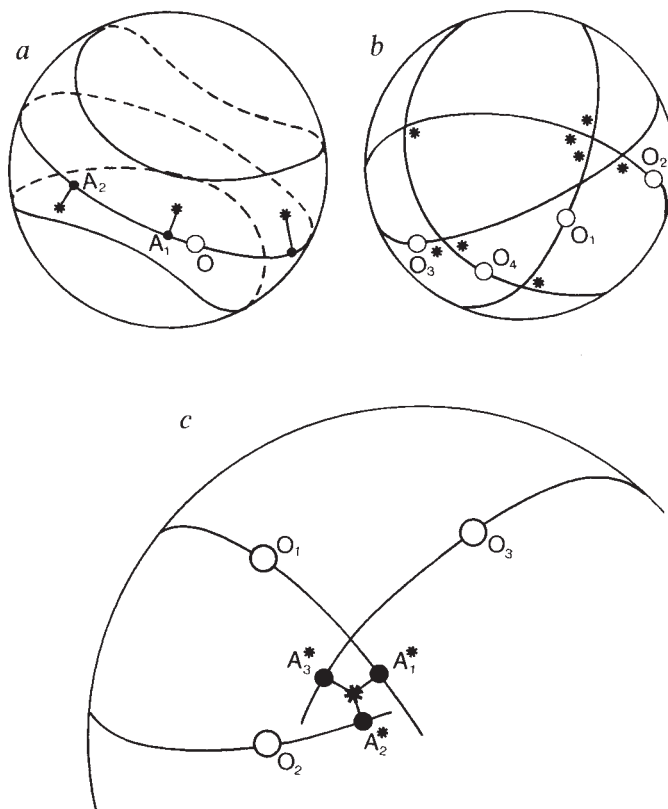


FIG. 3 The three steps of the scientific reduction of the data, illustrated with reference to the observations carried out during the mission. In (a), the star abscissae along 'reference great circles' are computed, based on the relative phase measurements of the modulated star signals. In (b), the origins of the complete net of some 2,000 reference great circles, accumulated throughout the mission, and which intersect at many different angles, are determined. In (c), the astrometric parameters of each star are extracted from the measurements on the basis of the star motions with respect to the resulting reference system.

■ The relative abscissae determined in the first step are converted into absolute abscissae by referring them to the origins of the system of reference great circles determined in the second step. Finally, the astrometric parameters of each star are determined by a least-squares combination of its abscissae on different reference great circles.

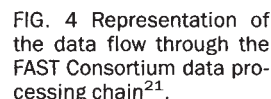
Calibration, photometry and double star analysis will be carried out in parallel with this three-step process. The accurate photometric measurements (to ~ 0.01 mag in m_{H} at ~ 8 mag for the 100 or so grid crossings each star will make during the mission) and the detailed detection of double or multiple stars with separations greater than ~ 0.1 – 0.2 arcsec for magnitude differences less than ~ 3 , will be important by-products of Hipparcos.

The attitude data will be used both for the on-ground attitude reconstruction and as input for the reductions leading to the compilation of the Tycho Catalogue¹⁵, comprising astrometric data of lower precision and two-colour photometric data for some 400,000 stars down to a limiting magnitude at ~ 11 . This project will rely on a separate input catalogue provided by a collaboration with the HST Guide Star Catalogue team.

In the design of the satellite, the development of the attitude and orbit control systems has had an important influence on performance. Because the basic measurements are essentially simultaneous comparisons, significant attitude jitter would have been tolerable. But it emerges that the attitude control system is smooth enough for stars not simultaneously present in the combined fields of view to still be tied to each other by taking account of the motion of the attitude of the satellite. This leads to an effective enlargement of the field of view, an improvement in the great-circle rigidity and thus to an improvement in the final star positions (in particular for the brighter stars which contribute directly to the rigidity of the reference frame).

The operation of the satellite in its geostationary orbit will be controlled from the ESA's operations control centre, ESOC, at Darmstadt, West Germany, which will continuously provide the characteristics of the programme stars to be observed. That information will determine both the observing programme and the scanning motion of the satellite. Data sent down from the satellite will be monitored, merged with auxiliary parameters such as information about the orbit of the satellite and despatched to the data reduction consortia for processing¹⁶.

Once on station, the satellite will be despun, its spin axis oriented towards the Sun and the 'scanning law' progressively acquired by the identification on the ground of star patterns with data from the star mapper. Payload commissioning is expected to take about 30 days, when preliminary estimates will be made of the geometrical and photometric properties of the payload and its stability in time. Quantities such as the modulation coefficients, detector response profile and star-mapper calibra-



tion parameters will also be determined. Routine operations will begin when the commissioning has been completed successfully.

Monitoring checks of the payload will be carried out at the ESOC control centre throughout the mission to quickly identify anomalies in payload operations. The activities of the data-reduction consortia, whose treatment of data is expected to lag 2–3 weeks behind its production, will also identify significant instrumental changes. But there is also a 'first-look' facility, implemented by the FAST consortium, that will process a set of about 12 hours of data each week to validate, as swiftly as possible, the data coming from ESOC and to provide the results of initial data calibration to both the FAST main processing chain and to the operations centre.

Stars to measure

Although the limiting visual magnitude of the observations is expected to be at $V \approx 12.4$, there will be not be enough time to gather complete information about all stars brighter than the limit, so that a selection of programme stars has been made. Accordingly, almost all 55,000 stars brighter than a completeness limit between 7 and 9 mag (depending on spectral type and galactic latitude) will be observed together with about 65,000 fainter stars between the completeness limit and the limit of observability.

The observing programme is based on a catalogue constructed by the Input Catalogue Consortium¹⁷. This input catalogue grew out of some 200 proposals received by ESA in 1982, involving some 214,000 stars, which were assigned scientific priority by a selection committee. Allowing for the technical limitations of the satellite, the possibility that data on some stars may be incomplete and inaccuracies or errors arise in the available ground-based data, the consortium has used detailed simulations to optimize the distribution of programme stars on the sky and to allocate observing times. As a result, the observation programme will be fixed from the outset of the mission.

To ensure that the *a priori* data are of sufficient quality, the consortium has complemented available data with cross-identifications, extensive ground-based astrometric and photometric observations and detailed work on particular categories of objects. Particular attention has had to be paid to double and multiple stars (where extensive re-measurement and re-cata-

TABLE 3 Composition of input catalogue

m_v	No. of stars	Priority 1 stars retained in catalogue (%)	Stars brighter than completeness limit retained in catalogue (%)
<6	4,400	99	99.3
6–7	8,800	99	99.3
7–8	23,100	98	93
8–9	43,100	97	91
9–10	27,600	90	...
10–11	8,850	85	...
11–12	2,600	85	...
>12	550	61	...
Total	119,000	...	...

Priority 1 stars are those allocated highest scientific priority by the scientific selection committee.

loguing has been necessary), minor planets (important for a definition of the dynamical reference system), large-amplitude variable stars (where a more precise knowledge of their ephemerides has been needed to allocate observing time), photometric standard stars and objects for establishing links to the extragalactic reference frame^{18–20}.

The following features of the work of the Input Catalogue Consortium deserve particular attention:

■ The definition of the completeness limit has involved considerations of the availability of observing time (and the consequent competition between stars brighter than the completeness limit and fainter, but astrophysically or astrometrically interesting, stars), the problem of statistical significance at the completeness limit (especially for variable stars) and the desirability of a uniform density of stars over the celestial sphere. The consortium has sought to optimize the value of the observations by minimizing the contribution of red giants, chiefly by means of a two-component definition of the limit involving spectral type and galactic latitude, b . (Specifically, for stars of spectral type earlier than or equal to G5, and for stars of other spectral types, the completeness limit is defined as $V_{\text{lim}} = 7.9 + 1.1 \sin b$ and $V_{\text{lim}} = 7.3 + 1.1 \sin b$, respectively. For stars of unknown spectral type, the first criterion is replaced by the second at the colour index $B - V = 0.8$ magnitude.) A few hundred stars known to be brighter than the completeness limit have been omitted from the input catalogue in densely populated areas of the sky.

■ Particular attention has been paid to certain key scientific programmes, for example the study of OB stars of different spectral types at different distances as well as stars contributing to the stellar luminosity function. Nearby stars and those significant in proper motion surveys are thus well represented, whereas there is a good distribution of Cepheid periods and amplitudes: the catalogue includes all the 55 Cepheids nearer than 1,000 pc, all but one of the objects used for the calibration of the period-luminosity-colour relation and all 26 RR Lyrae stars within 1,000 pc. The choice of stars in galactic clusters and the Magellanic Clouds has been carried out by special working groups. As an example, about 100 stars from the Hyades Cluster and a further 100 stars at its edges are included. There is a good representation of spectroscopically peculiar stellar types within 400 pc, and all known metal-poor stars within 20 pc are included. Of the 45 proposed white dwarfs, 21 will be observed; the others have been rejected either because proximity to bright stars would perturb the expected signals, or because they are too faint.

■ Among stars of predominantly astrometric interest, the catalogue includes 96 per cent of those listed as International Reference Stars (38,300 stars out of 40,500) and all FK4 (the predecessor of the FK5) stars. Proposals to ESA concerned with linking the optical and radio reference frames included 186 radio stars, of which 183, directly observable by Hipparcos and by

Hipparcos collaborators

ESA Project Manager: H. Hassan; ESA Project Scientist: M. A. C. Perryman; MATRA Project Manager: M. Bouffard; Aeritalia Project Manager: B. Strim.

Past and present consortium task leaders:

FAST Consortium: J. Kovalevsky (Team Leader), P. L. Bernacca, D. T. van Daalen, F. Donati, M. Froeschlé, I. Galligani, D. Iorio-Fili, H. van der Marel, F. Mignard, B. Morando, F. P. Murgolo, J. L. Pieplu, R. A. Preston, S. Röser, H. Schrijver, H. G. Walter.

INCA Consortium: C. Turon (Team Leader), A. N. Argue, A. Bec-Borsenberger, M. Crézé, J. Dommanget, D. Egret, A. Gómez, M. Grenon, H. Jahreis, T. Lederle, M. O. Mennessier, J. C. Mermilliod, L. Prévot, Y. Réquière.

NDAC Consortium: E. Høg (Team Leader), A. M. Cruise, F. van Leeuwen, L. Lindegren, N. Lund, C. A. Murray, M. J. Penston, C. Petersen, N. Ramamani, M. A. J. Snijders, S. Söderhjelm.

TDAC Consortium: E. Høg (Team Leader), G. Bässgen, U. Bastian, P. L. Bernacca, D. Egret, M. Grewing, J. L. Halbwachs, J. Kovalevsky.

Chairman, Scientific Selection Committee: A. Blaauw. These names represent some 100 scientific personnel involved in the preparation of the mission and data treatment, some 40 ESA staff and some 1,800 industrial personnel throughout those ESA member states involved in the development of the satellite.

TABLE 4 Standard errors (in milliarcsec) of the five astrometric parameters as function of stellar magnitude

	<6.5	6.5–7.5	7.5–8.5	8.5–9.5	9.5–10.5	10.5–11.5	>11.5
m_H	<6.5	6.5–7.5	7.5–8.5	8.5–9.5	9.5–10.5	10.5–11.5	>11.5
$\sigma_{\alpha \cos \delta}$	1.04	1.06	1.13	1.24	1.41	1.84	2.37
σ_{δ}	0.87	0.90	0.96	1.03	1.17	1.51	1.96
$\sigma_{\mu \alpha \cos \delta}$	1.47	1.52	1.61	1.76	2.00	2.62	3.37
$\sigma_{\mu \delta}$	1.21	1.25	1.32	1.43	1.62	2.08	2.72
σ_{π}	1.26	1.30	1.39	1.52	1.74	2.26	2.94

The values are sky averages for stars in different magnitude intervals. The figures represent the most recent estimates of the final catalogue errors made by the data reduction consortia (L. Lindegren, personal communication).

TABLE 5 Variation of astrometric standard errors with ecliptic latitude

$ \beta $	0–6°	6–12°	12–17°	17–24°	24–30°	30–37°	37–44°	44–53°	53–64°	>64°
α	1.32	1.30	1.27	1.22	1.14	1.02	0.85	0.66	0.73	0.77
δ	1.10	1.09	1.09	1.08	1.06	1.02	0.96	0.89	0.89	0.91
π	1.19	1.19	1.18	1.15	1.11	1.06	0.98	0.86	0.76	0.69

The table gives the factor to be applied on the sky averages in Table 4. The same factor can be used for the proper motions (μ_{α} and μ_{δ}) as for the positional components (α and δ).

radio VLBI, remain in the catalogue. Of the list of 174 stars proposed as links with observations by the Hubble Space Telescope, all but one will be observed by Hipparcos. These stars lie near quasars. The ultimate goal of these observations is that the final Hipparcos reference system will itself be quasi-inertial and the proper motions absolute. Some forty minor planets are also candidates for observation; the ground-based observations needed to provide their positions to sufficient accuracy are continuing.

The distribution of stars in the final catalogue is given in Table 3 as a function of 'Hipparcos magnitude', m_H . (The relationship between m_H and the Johnson B and V magnitudes is of the form: $m_H = m_V + 0.38(B-V) - 0.13(B-V)^2$.) The input catalogue — the complete list of stars in the final Hipparcos Catalogue together with all available information about them — is expected to be published in early 1991.

Making a catalogue

Averaged over the sky, the expected precision of the two components of position, the annual proper motion and the parallax, is ~ 0.002 arcsec for stars brighter than ~ 9 mag, degrading to ~ 0.004 arcsec for stars at the limit of observability at ~ 12 mag. Actual precision will depend weakly on ecliptic latitude, star colour, local stellar density and sky background (see Tables 4,5). If the satellite performs and is operated as planned, a first full-sky catalogue of improved star positions will be available after about six months of observations. As time passes, decoupling star positions from parallaxes and proper motions will slowly become more certain²¹.

The treatment of the 2.5 years of data is a major undertaking, and has been recognized from the outset to require a considerable amount of computing effort and analysis. Two separate consortia, NDAC and FAST, are responsible for the main task of data reduction, for which software development began in 1982 by means of collaboration between ESA member-state institutes. Both consortia have followed the three-step procedure described above, but the algorithms and detailed procedures they have chosen are sometimes quite different. This is not

surprising, for the number of executable FORTRAN statements or system procedures involved in the complete reduction software ranges between 100,000 and 150,000. Each consortium has spent much effort on the preparation of simulated data, chiefly to test the programs developed without waiting for the satellite to be commissioned, but the exchange of simulated data and comparisons of the results obtained have allowed both consortia to check and sometimes to correct their programs.

The essential reason why there are two groups is that the ultimate users of the unified Hipparcos catalogue can be confident of its reliability. But it would evidently be undesirable to have to wait for two data-reduction teams to complete separate catalogues before attempting to compare them. Comparisons of the outputs of the various reduction steps using simulated data have already been made, and will continue as real data begin to flow. The intention is to identify where in the two sets of programs different results appear, to investigate the reasons for this divergence and, it is hoped, to identify and correct errors in the algorithms used or in their execution. The many differences in the representation and treatment of the two consortia should strengthen confidence in the final results, but they complicate considerably inter-consortium comparison.

There is a third international data-reduction consortium, TDAC, to compute the astrometric and photometric results from the data acquired by the star mapper (the Tycho data). Formed in 1983, the consortium has again relied on several international teams to prepare the software, which has been verified by the use of simulated data. But in this context, it was considered neither possible nor necessary that there should be two independent consortia working in parallel.

For an assumed mission duration of 2.5 years (consumables on board the satellite are sized for a somewhat longer duration) the final Hipparcos and Tycho Catalogues should be available to the astronomical community by 1995. $\square$

M. A. C. Perryman is Project Scientist for Hipparcos, at ESA's technology centre: ESTEC, Postbus 299, 2200 AG Noordwijk zh, The Netherlands.

1. Lacroute, P. in *Proc. 13th IAU Gen. Assembly* (Prague, 1967).
2. Argue, A.N. in *IAU Highlights of Astronomy* (ed. Swings, J.-P.) **7**, 719–723 (1986).
3. Lestrade, J.-F. *et al. Astr. J.* **96**, 1746–1752 (1988).
4. Monet, D.G. *A. Rev. astr. Astrophys.* **26**, 413–440 (1988).
5. Kovalevsky, J. *Space Sci. Rev.* **39**, 1–37 (1984).
6. Bernacca, P.L. *Astrophys. Space Sci.* **110**, 21–45 (1985).
7. Schuyler, M. *ESA Bull.* **42**, 22–29 (1985).
8. Bouffard, M. in *19th Eur. Space Symp.* (NASA Goddard SFC, 1985).
9. *IAU Highlights of Astronomy* (ed. Swings, J.-P.) **7** (1986).
10. Lindegren, L. in *Astrometric Techniques* (eds Eichhorn, H.K. & Leacock, R.J.) *IAU Symp.* **109**, 593–599 (Reidel, Dordrecht, 1986).
11. Bernacca, P.L. (ed.) *Proc. FAST Thinkshop, Asiago* (1983).
12. Kovalevsky, J. (ed.) *Proc. Second FAST Thinkshop, Marseilles* (1985).
13. Bernacca, P.L. & Kovalevsky, J. (eds) *Proc. Third FAST Thinkshop, Bari* (1986).

14. *Manuscripta Geodetica* **11** (No. 2) (Springer, Berlin, 1986).
15. Hög, E. *IAU Symp.* **109** (1985).
16. van der Ha, J. *ESA Bull.* **51**, 27–33 (1987).
17. Turon, C. in *IAU Symp.* (eds Débarbat, S. *et al.*) **133** (Reidel, Dordrecht 1988).
18. *Proc. Aussois Colloq. Scientific Aspects of the Input Catalogue Preparation* (eds Turon, C. & Perryman, M.A.C.) *ESA SP234* (1985).
19. *Proc. Sitges Colloq. Scientific Aspects of the Input Catalogue Preparation* (eds Torra, J. & Turon, C.) *CIRIT* (1988).
20. *Proc. Strasbourg Colloq. Scientific Aspects of the Hipparcos Mission* (eds Perryman, M.A.C. & Guyenne, T.D.) *ESA SP177* (1982).
21. *The Hipparcos Mission: Pre-Launch Status* (eds Perryman, M.A.C. *et al.*) *ESA SP1111* (1989).
22. Lindegren, L. & Kovalevsky, J. in *IAU Highlights of Astronomy* (ed. Swings, J.-P.) **7**, 699–706 (1986).

Stratospheric clouds and ozone depletion in the Arctic during January 1989

D. J. Hofmann*, T. L. Deshler*, P. Amedieu†, W. A. Matthews‡, P. V. Johnston‡, Y. Kondo§, W. R. Sheldon||, G. J. Byrne|| & J. R. Benbrook||

* Department of Physics and Astronomy, University of Wyoming, Laramie, Wyoming 82071, USA

† Service d'Aéronomie, CNRS, 91371 Verrières le Buisson, France

‡ DSIR/PEL Lauder, Central Otago, New Zealand

§ Research Institute of Atmospheric Sciences, Nagoya University, Toyokawa, Japan

|| Department of Physics, University of Houston, Houston, Texas 77204, USA

Stratospheric clouds, believed to be necessary for springtime polar ozone depletion to take place, were detected with balloon-borne sensors at Kiruna, Sweden during January 1989, the coldest January in the north polar stratosphere for at least 25 years. Comparison of the ozone profile in the region of the clouds with those obtained during the past three austral spring seasons at McMurdo Station in Antarctica suggests the beginning of ozone depletion at a height of 22–26 km.

THE role of stratospheric clouds in springtime polar ozone depletion has been recognized in the Antarctic but not, so far, in the Arctic. Here we describe measurements made from two instrumented balloons in January 1989 with which we have been able to establish both the appearance of Arctic stratospheric clouds and the signature of the onset of ozone depletion at or around the altitude of these clouds.

The Arctic polar stratosphere during the winter of 1988–89 was unusually cold with no stratospheric sudden warmings until February¹. The polar vortex was relatively stable throughout January, giving rise to conditions similar to those in the Antarctic winter stratosphere. The January average North Pole 30-hPa (~22-km altitude) height and temperature were the lowest ever measured. The average temperature at 30 hPa for January was as much as 12 °C below the 20-year mean over large areas of the polar region. Figure 1 shows the 30-hPa heights and temperatures on 23 and 24 January and on 30 and 31 January, periods covering the two balloon flights to be discussed here. Before 23 January, the vortex was centred near the North Pole. Intensification of the Aleutian high-pressure system began a shift of the vortex towards Scandinavia on 24 January which was eventually followed by a major stratospheric warming on about 10 February (K. Labitzke, personal communication). Before the major warming, however, temperatures as low as –92 °C were observed on 1 February at 30 hPa over southern Scandinavia. Temperatures below –85 °C in the 20–25-km region were responsible for the frequent displays of nacreous or 'mother of pearl' clouds (stratospheric water-ice crystals) observed at Kiruna, Sweden (68 °N, 20 °E).

Stratospheric clouds are thought to be important in converting hydrochloric acid and chlorine nitrate into active chlorine and in stratospheric denitrification through the condensation of HNO₃ vapour, thus removing odd nitrogen from the gas phase which would otherwise prevent the destruction of ozone by the active chlorine². The observation of extensive stratospheric clouds in the Arctic in January 1989 is therefore important in assessing ozone depletion under perhaps the most favourable conditions for this to occur since at least 1957. Before 1957 there was insufficient chlorine in the stratosphere to cause ozone depletion even with extensive clouds.

At a pressure of 30 hPa and for a typical stratospheric water-vapour mixing ratio of 5 p.p.m.v., a temperature of ~–87.5 °C is required to condense the vapour and form ice clouds (the 'frost point' temperature). However, studies of the nitric acid trihydrate (NAT) solid in the laboratory³ indicate that this crystalline phase will form at a temperature of –80 °C (10 parts per 10⁹ by volume (p.p.b.v.) HNO₃) to –81 °C (5 p.p.b.v. HNO₃) for 5 p.p.m.v. water vapour at 30 hPa. A predominant NAT composition of Antarctic stratospheric clouds was proposed earlier from theoretical considerations^{4,5} and has been identified in *in situ* measurements in Antarctica⁶. Thus, it appears that extensive NAT clouds and more local nacreous (water-ice) clouds should have formed in the Arctic stratosphere above ~20 km during January 1989. Because sunlight is also required to destroy ozone, the question of Arctic ozone depletion hinges on the amount of time that the air, which has been heterogeneously processed by stratospheric clouds, has spent in sunlight. As the Arctic polar vortex generally breaks down before sufficient sunlight reaches the near-polar regions, the most probable location for the detection of ozone depletion in the Arctic is near the Arctic circle during unusually cold years. For example, Kiruna experienced ~8 hours of sunlight per day on 23 January at 21 km where temperatures as low as –86.7 °C were observed. Owing to the effects of the Aleutian high, the vortex and the coldest regions are generally displaced towards Scandinavia (see Fig. 1) so that locations such as Kiruna are quite favourable. Here we report observations of polar stratospheric clouds and ozone during balloon flights on 23 and 30 January at Esrange, near Kiruna.

Cloud particle measurements

The data were obtained with University of Wyoming balloon-borne particle counters as used on many occasions in Antarctica⁷. The optical counter has an air-sample flow rate of ~200 cm³ s^{–1}, capable of resolving low concentrations (~10^{–3} cm^{–3}) as associated with ice-crystal formation. The instrument measured in seven size ranges, radius $r \geq 0.20, 0.25, 0.30, 1.0, 2.0, 3.0$ and $5.0 \mu\text{m}$, for spherical particles or slightly non-spherical particles of equivalent volume. A condensation-nuclei detector sensitive to particles with $r \geq 0.01 \mu\text{m}$ was added for the second flight on 30 January to obtain the total aerosol concentration profile and to determine the fraction of condensation sites taking part in cloud-particle growth.

Figure 2 shows the temperature and $r \geq 0.20 \mu\text{m}$ aerosol profiles measured during balloon ascent for the two flights. For the flight of 23 January, temperatures were <–80 °C between altitudes of 18.5 and 27 km and below –85 °C between 20.2 and 21.3 km, and reached –85 °C briefly several times between 23 and 25 km. Because the balloons encountered very low temperatures, the rise rate was generally kept low (1–2 m s^{–1}) with occasional short periods of float before ballasting. This may have resulted in inadequate ventilation of the temperature sensor and could account for some of the unusual temperature increases observed above ~20 km. This would not affect the particle

counter or ozone sensors because they are aspirated by pumps. Some of the temperature structure seen during both flights may be related to orographic waves originating from the mountain ranges west of Kiruna which appeared to be responsible for the generation of nacreous clouds during the measurement period. Westerly wind speeds in the stratosphere were very high (in excess of 50 m s^{-1}) and a minimum temperature during the balloon ascent on 23 January of -86.7°C was reached in the cold layer between 20 and 21 km. The second flight, on 30 January, did not experience such low temperatures, with a minimum of -82.7°C at $\sim 25 \text{ km}$.

The aerosol profiles in Fig. 2 indicate a normal sulphate layer between ~ 10 and 18 km with concentrations similar to those observed in Antarctica in September 1988⁸. These aerosols represent the global background sulphuric acid layer, there having been no major volcanic eruptions since El Chichón in 1982 which increased the atmospheric sulphate concentration by nearly an order of magnitude⁹. In the cold layer at 19–22 km on 23 January we see a large increase in $r \geq 0.20 \mu\text{m}$ particles, reaching a concentration of 8 cm^{-3} or more than 50% of the

concentration of condensation nuclei in this region. The condensation-nuclei or total-aerosol profile was measured on the second flight and is reasonably constant in the absence of homogeneous nucleation of new particles. The particle enhancement on 23 January actually extends uniformly at lower concentrations to at least 26 km.

The large particles ($r \geq 1.0 \mu\text{m}$) were also enhanced but only to the extent of ~ 1 in 10^4 of the available condensation nuclei. Very few particles having radii $> 2 \mu\text{m}$ and no large ice crystals ($r \geq 5 \mu\text{m}$) were observed in the layer. Thus it appears that pure water did not condense in this cloud, indicating a frost point temperature below -87°C at 36 hPa or a water-vapour mixing ratio of $\leq 4.5 \text{ p.p.m.v.}$ No large ice crystals were detected at any other altitude and the balloon therefore did not penetrate any nacreous clouds which may have been present at the time. Overcast skies prevented the observation of such clouds on the day of the flight. Stratospheric clouds, however, were visible nearly every day when observing conditions allowed, on some occasions to solar zenith angles of 96° , indicating altitudes in excess of 20 km, throughout the period 23 January to 2 February.

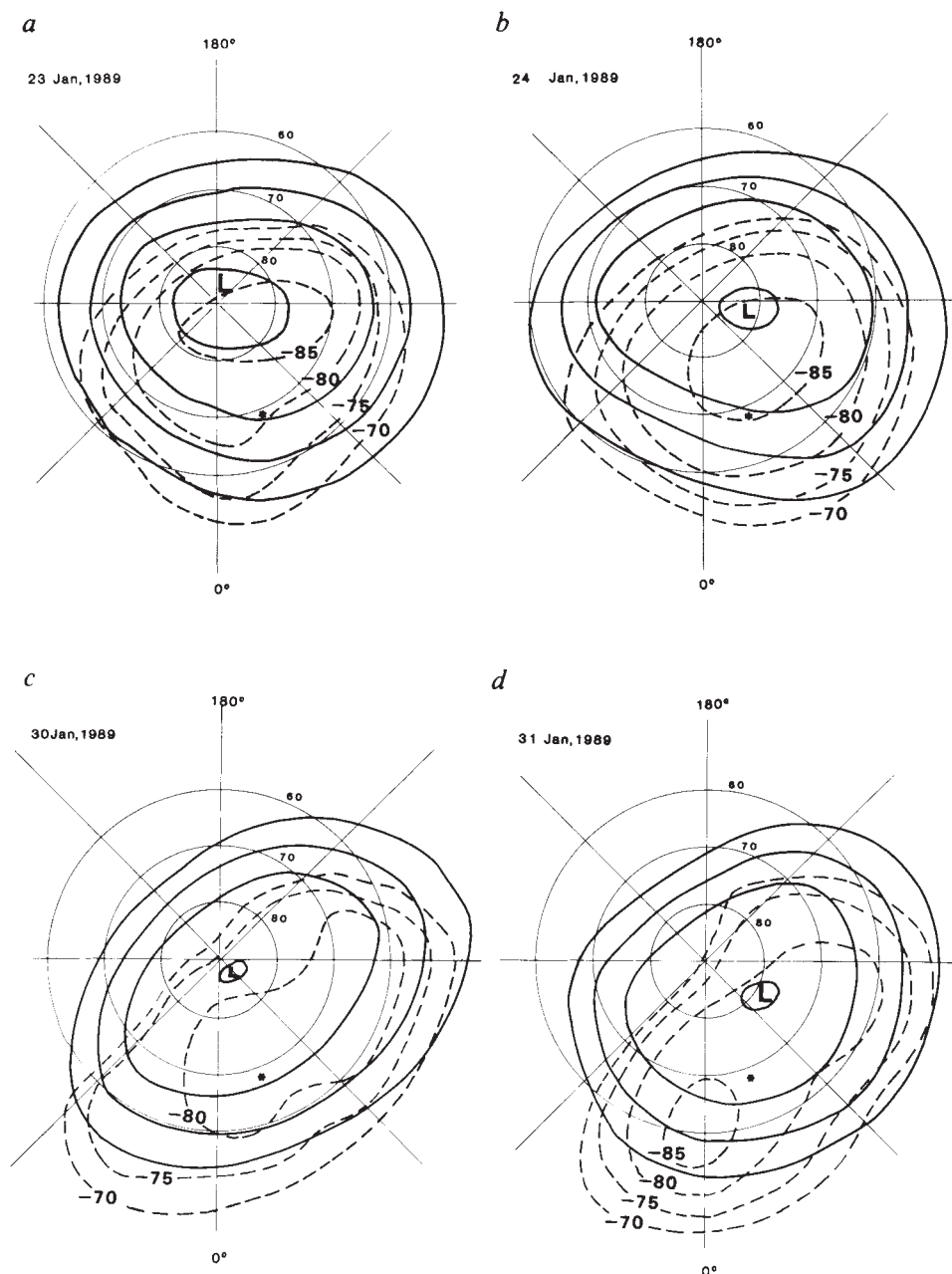


FIG. 1 Temperature (dashed) and height (full) contours at 30 hPa in the Arctic for 00:00 GMT on 23 January (a), 24 (b), 30 (c) and 31 (d), 1989, from the Berlin analysis (K. Labitzke, personal communication). The height contour of the north polar low is 21.5 km and the contour interval is 0.5 km. Temperatures are in $^\circ\text{C}$. The location of Kiruna is marked by a star. Balloon flights at Kiruna occurred from 14:30 to 17:30 GMT, 23 January, and 11:30 to 14:30 GMT, 30 January.

Very low values of column NO_2 were measured by the PEL (Physics and Engineering Laboratory, Lauder, New Zealand) group at the Swedish Institute of Space Physics Observatory in Kiruna during the measurement period. These low values are essential for the chlorine chemistry detailed in ref. 3 to occur. The second flight on 30 January, at somewhat warmer temperatures, observed numerous thin (100–300-m-thick) particle layers between 18 and 26 km. In contrast to the 23 January cloud, the particles in these thin layers were mainly in the 2–3- μm radius size range at concentrations of $<10^{-2} \text{ cm}^{-3}$ suggesting preferential growth on the largest ($r \geq 0.30 \mu\text{m}$) sulphate aerosol as observed previously in Antarctica⁷. It is quite likely that the 18–26-km region contained such tenuous layers of large cloud particles throughout the period 23–30 January.

Ozone observations

Ozone profiles were measured with the Service d'Aeronomie chemiluminescent detector¹⁰ on the 23 January flight, the University of Houston (NASA-JSC) ultraviolet photometer¹¹ on the 30 January flight and the University of Wyoming digital electrochemical detector¹² on both flights. Figure 3 shows the ozone partial pressure and mixing-ratio profiles measured by these instruments during the two flights. Agreement between the different techniques on both flights is generally good, although the electrochemical instrument is known to have uncertainties owing to the pump efficiency, especially at high altitude¹³. The pumps on these instruments were calibrated before flight. In addition, the response time is longer (20–30 s) for the electrochemical sensor as compared to the other two. This may explain slight displacements of some of the peaks, especially on 30 January when the balloon rise rate was considerably faster. In addition, each ozone sensor had its own ambient pressure sensor, and some small differences may therefore be expected. Integration of the partial pressure profiles resulted in ~ 325 Dobson units (DU) of total column ozone on both days. Column measurements of total ozone by the PEL group at Kiruna indicate a value of 330 ± 25 DU on 23 January. The total ozone mapping spectrometer (TOMS) satellite instrument gave values of ~ 270 DU and 235 DU near Kiruna on the two days (A. Krueger, personal communication). 23 January was the first day for which TOMS had sufficient sunlight to determine a value. Past experience at McMurdo in Antarctica indicates that TOMS values are $\sim 10\%$ lower than those calculated from ozonesonde data during the first days of sunlight in early September but then generally agree to within a few per cent thereafter¹³. The large fluctuations in ozone between 12 and 18 km are similar to those observed in association with quasi-horizontal advection, possibly from other latitudes¹⁴. However, the reductions in ozone in the 22–26-km range of $\sim 25\%$ on 23 January and $\sim 10\%$ on

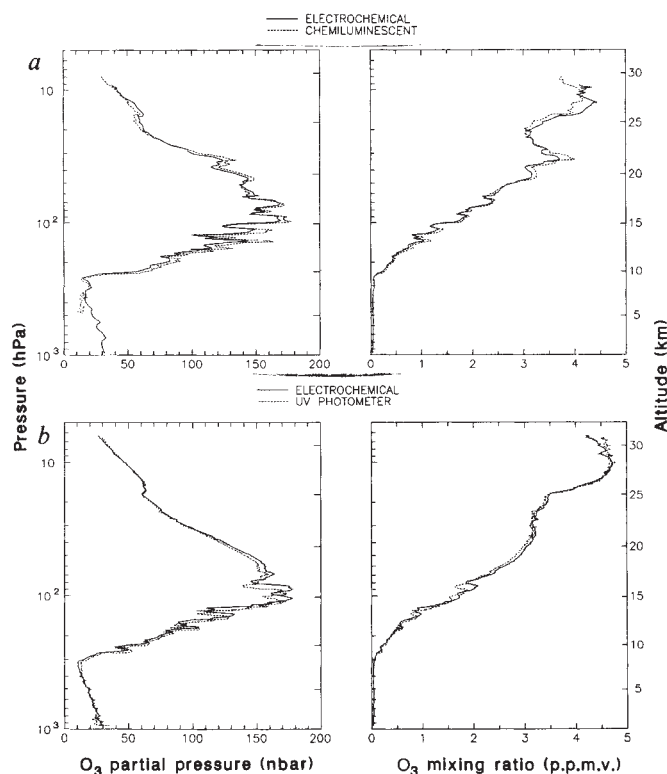


FIG. 3 Ozone partial pressure and mixing ratio profiles measured at Kiruna, Sweden on 23 January (a) and 30 January (b), 1989.

30 January seem unusual. As discussed below, this apparent ozone deficit is very similar to the initial stages of ozone depletion in early September in Antarctica.

Discussion

The question of denitrification, necessary for ozone depletion to take place, can be addressed from cloud particle observations on the assumption that they are composed of NAT. Conversion of the particle-size distribution to a mass concentration, or the determination of the condensation temperature, can be used to estimate the amount of HNO_3 vapour present in the cloud.

Some information on the condensation temperature may be obtained by comparing data during balloon ascent and parachute descent, which occurred ~ 1 h after passing through the cloud at 21 km on 23 January. Winds were from the north-west at a relatively constant speed of 50 m s^{-1} above 20 km so that the cloud observed on ascent probably should have also been observed on descent if it persisted. The ascent and parachute descent profiles of temperature and aerosol concentration are compared in Fig. 4. The descent temperature profile has less structure because of the rapid descent rate in the stratosphere. Small fluctuations in the sulphate layer at 14 km, reproduced in the descent profile, further suggest little differential motion of the balloon and the 15–20-km stratospheric air mass. The descent profile indicates the lowest temperature to be $\sim -87^\circ\text{C}$ at 24 km. It should be noted that the temperature sensor on the particle counter, mounted on top of the gondola, was not ideal for measuring temperatures during the initial rapid descent. A side-mounted thermistor of the Nagoya group, which gave similar results on ascent, gave a temperature of $\sim -90^\circ\text{C}$ at 24 km. In this very cold region another particle layer, not present on ascent, was observed. The 19–22-km particle layer, observed on ascent at a concentration of $\sim 8 \text{ cm}^{-3}$, was again present at a temperature of $\sim -84^\circ\text{C}$ (on both sensors) and a reduced concentration of $2\text{--}3 \text{ cm}^{-3}$. One can probably conclude that

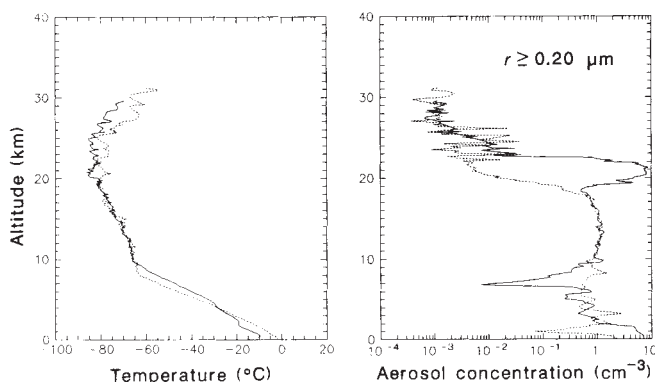


FIG. 2 Temperature and aerosol ($r \geq 0.20 \mu\text{m}$) profiles measured during balloon ascent at Kiruna, Sweden on 23 January (full lines) and 30 January (dashed lines) 1989.

some evaporation of the cloud observed on ascent had occurred at the higher temperature and that a new cloud had formed or was forming at 24–25 km, suggesting a condensation–evaporation temperature of $\sim -84^\circ\text{C}$.

A condensation temperature of -84°C for NAT is somewhat lower than expected and suggests either substantial denitrification had already taken place in this air parcel or that it was undergoing rapid cooling, in which case the vapour may have been supersaturated. For 5 p.p.m.v. water vapour at 36 hPa (the centre of the cloud) and a NAT condensation temperature of -84°C , laboratory measurements³ suggest a HNO_3 mixing ratio of ~ 0.25 p.p.b.v. or nearly complete denitrification. On the other hand, if the actual HNO_3 mixing ratio was as high as 10 p.p.b.v., as measured in the 20–25-km region at Kiruna on 1 February 1988 (ref. 15), a very high degree of supersaturation would have had to exist. It is quite likely that neither of these conclusions is correct by itself but that this air parcel had been somewhat denitrified and that the cold layer at 21 km was caused by a recent rapid cooling.

To address this issue further, we may estimate the mass in the cloud-particle distribution. The measured integral size distributions can be fitted with differential log-normal distributions to facilitate calculation of mass and particle surface-area density⁷. The aerosol mass can be converted to a condensed vapour mass for an assumed NAT composition. Table 1 summarizes these calculations for the small and large particle mode in the 21-km layer on 23 January. Mass estimates for the thin layers of large particles observed on 30 January are not included here because of uncertainties in the size owing to the unknown shape of these larger particles. However, if composed of NAT and if the 2–3- μm -radius estimate is even approximately correct, then some degree of denitrification through sedimentation could have occurred. Further cooling below the water frost point for only a short period could result in considerable denitrification without substantial dehydration¹⁶. Size determination, and therefore estimation of mass, is much more accurate for the small particle mode which comprised most of the mass in the 23 January layer. The implied HNO_3 mass of ~ 5 p.p.b.v. incorporated in the aerosol suggests that if the cloud observed on 23 January was indeed composed of NAT, then the air in the cloud was probably not completely denitrified. Although the small particles are inefficient for denitrification, as indicated in Table 1, they dominate the surface-area density, the parameter important for heterogeneous chemistry. This was also observed in Antarctic polar stratospheric clouds at lower altitude⁸.

We now turn to the issue of possible ozone depletion in the 22–26-km region, where the mixing ratio was seen to be depressed (see Fig. 3). Transport of ozone-poor air from other latitudes is not expected here owing to the strength and containment properties of the vortex at these altitudes. One cannot

entirely rule out the possibility that the depression is a permanent feature of the Arctic ozone distribution; however, similar depressions are not observed in Antarctica before the beginning of ozone depletion in late August. The 30 January ozone reduction might be interpreted as having been due to a major vertical motion between about 20 and 26 km, because the ozone mixing ratio is relatively constant over this region; however, the temperature profile does not appear to vary adiabatically over this interval as might be expected. On the other hand, the reduction in ozone mixing ratio with altitude between 21 and 24 km on 23 January suggests a true ozone sink. The altitude of the ozone deficit does not coincide exactly with the large increase in small particles seen on 23 January but does coincide with the region of lower concentrations of larger, probably aged, particles as seen both on 23 and 30 January. The numerous small particles at 19 to 22 km on 23 January are believed to have been created

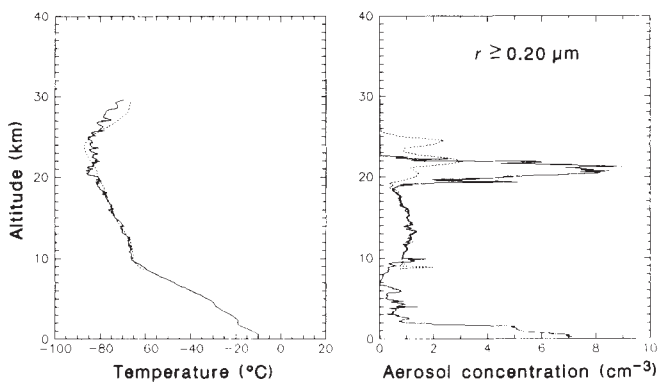


FIG. 4 Temperature and aerosol ($r \geq 0.20 \mu\text{m}$) profiles measured during balloon ascent (full lines) and parachute descent (dashed) at Kiruna, Sweden on 23 January 1989.

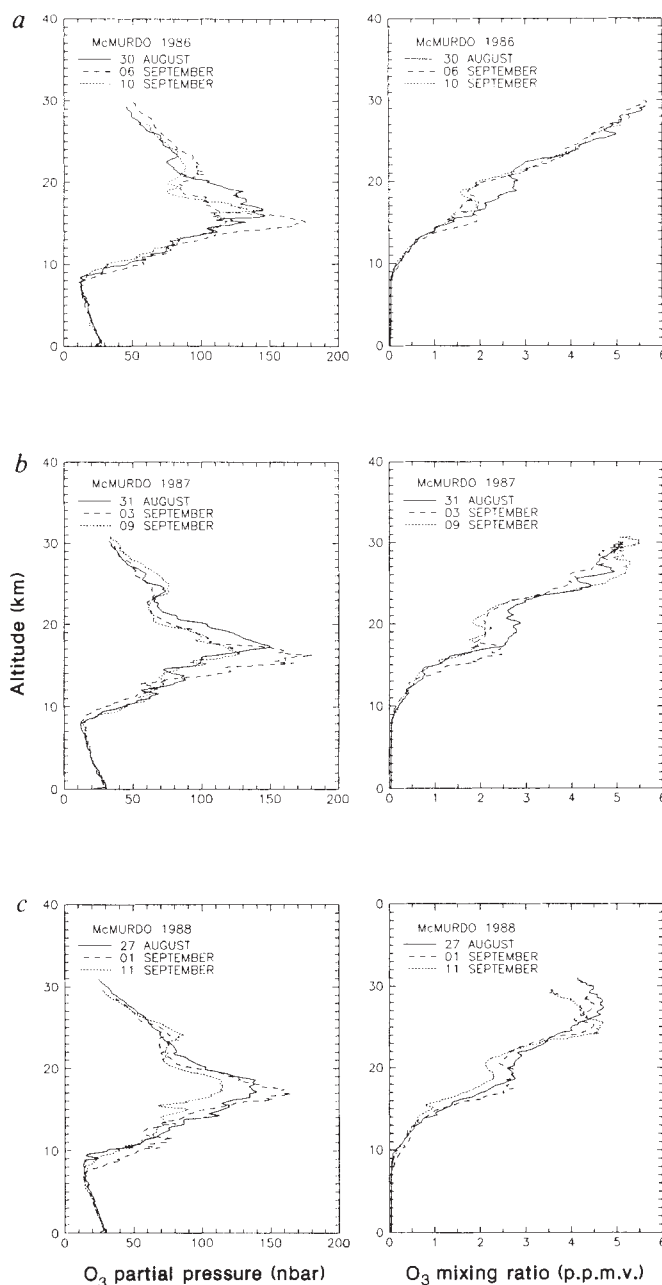


FIG. 5 Ozone partial pressure and mixing-ratio profiles measured at McMurdo Station, Antarctica during the early phases of springtime ozone depletion in 1986 (a), 1987 (b) and 1988 (c).

TABLE 1 Polar stratospheric cloud mass and area on 23 January at a height of 21 km

Particle size mode	N_0 (cm^{-3})	σ	r_0 (μm)	Mass† (p.p.b.v.)	Area ($\mu\text{m}^2 \text{cm}^{-3}$)
Small	15	1.36	0.20	4.7	9.1
Large	0.002	1.59	1.5	0.4	0.09
Total				5.1	9.2

N_0 , σ and r_0 are respectively the total number concentration, the distribution width and the mode radius of a log-normal size distribution.

† HNO_3 vapour mass from assumed nitric acid trihydrate composition, density = 1.62 g cm^{-3} .

recently through rapid cooling by vertical motions, possibly a lee-wave event, which resulted in the very low temperature in the 21-km region as discussed earlier. Thus, this cloud was probably not old enough to have a measurable effect on ozone chemistry.

As mentioned earlier, although stratospheric sunlight in late January is insufficient for the photochemistry required for ozone depletion over much of the Arctic, this is not the case at Kiruna, near the Arctic circle, especially if slightly non-zonal flow increases the time air parcels spend in sunlight, as was the case in late January 1989. For example, from the Berlin analysis of Northern hemisphere 30-hPa heights, winds and temperatures (K. Labitzke, personal communication), simple isobaric trajectory estimates suggest that an air parcel may have been in sunlight for as much as half of the time. The temperature extremes along this approximate trajectory are large owing to the fact that the height and temperature contours were not concentric (see Fig. 1) with variations from -57 to -86°C . It thus appears that the air sampled over Kiruna in late January may have been subjected to the required conditions for ozone depletion, that is, temperatures sufficiently low for denitrification through the formation of polar stratospheric clouds and also sufficient sunlight. However, as the timescale for the onset and relaxation of chemical ozone depletion may be the order of a week or more, such air-parcel history estimates may not be definitive in determining the ultimate extent of ozone depletion.

It has been shown¹²⁻¹⁴ that ozone depletion begins at McMurdo Station, Antarctica (78°S) in the 22-km region during the last week in August and then progresses rapidly to lower altitude, as stratospheric illumination rapidly increases, terminating in the 12-km region. There are some interesting similarities in the apparent deficit in ozone mixing ratio observed at high altitude in the Arctic and that observed in Antarctica. The 20-km region experienced ~ 8 h of sunlight per day at Kiruna

on 23 January and ~ 9 h on 30 January. Similar illumination periods exist at McMurdo in late August and early September. There are some slight differences at these times as the Sun rises several degrees higher in the sky at local noon at Kiruna than at McMurdo whereas the daily increase of sunlight hours is slightly faster at McMurdo than at Kiruna.

Figure 5 shows ozone partial pressure and mixing-ratio profiles covering this period at McMurdo in 1986, 1987 and 1988. McMurdo was in a favourable position well within the polar vortex during the last week of August and the first half of September in each of these years so that temperature histories above 20 km are comparable. The general nature of the ozone mixing-ratio depression is strikingly similar to that observed at Kiruna except that it is in the 18–22-km region, varying slightly from year to year, or ~ 4 km lower than in the Arctic. Thus, it is quite likely that we are witnessing the onset of a phenomenon identical to that which takes place in the Antarctic stratosphere except at higher altitude in the Arctic. In Antarctica the depletion generally continues throughout September and October, reaching altitudes as low as 12 km because these regions are very cold and clouds form throughout the winter, a situation which does not exist in the Arctic. For example, the 30-hPa average polar temperature in January 1989 was $2-3^\circ\text{C}$ colder than at 50 hPa (about 19 km altitude) and $\sim 8^\circ\text{C}$ colder than at 100 hPa (~ 15 km altitude). The latter temperatures were too high for cloud formation (-75°C minimum) whereas temperatures of -80 to -85°C are common at 15 km in the Antarctic winter and spring. Thus, the low temperatures required for ozone depletion are at higher altitude in the Arctic, where there is much less ozone than in the 12–22-km Antarctic depletion range.

The amount of the apparent ozone deficit observed in the Arctic is very small. If the mixing-ratio curve on 23 January is filled in so that the depression between 22 and 26 km does not exist, and the total ozone is then calculated, a value only 10 DU, or $\sim 3\%$, higher is obtained. In more normal, relatively cloud-free years, one would expect even smaller reductions. Thus, although 1989 was an unusually cold winter, the major warming and breakdown of the polar vortex in early February probably terminated any possibility for ozone depletion of sufficient magnitude for surface or satellite measurements to detect easily. The warming ended the cloud formation process and cloud evaporation reintroduces nitrogen to the gas phase, reducing any ozone depletion that may have been occurring. Some irreversible loss of odd nitrogen, through cloud-particle sedimentation during the winter, may have occurred. If the cold regions had persisted into February or March, a larger depletion could have resulted, but in any case could not have exceeded removal of most of the ozone between about 18 and 26 km, the cloud formation region, which is at most 25% of the total. □

Received 2 March; accepted 2 June 1989.

- Kuhlbarsh, T. & Naujokat, B. *Beilage zur Berliner Wetterkarte* 25/89 Institute for Meteorology, Free University of Berlin (1989).
- Solomon, S. *Rev. Geophys.* **26**, 131–148 (1988).
- Hanson, D. & Mauersberger, K. *Geophys. Res. Lett.* **15**, 855–858 (1988).
- Toon, O. B., Hamill, P., Turco, R. P. & Pinto, J. *Geophys. Res. Lett.* **13**, 1284–1287 (1986).
- Crutzen, P. J. & Arnold, F. *Nature* **324**, 651–655 (1986).
- Fahey, D. W. *et al.* *J. geophys. Res.* (in the press).
- Hofmann, D. J., Rosen, J. M., Harder, J. W. & Hereford, J. V. *J. geophys. Res.* (in the press).
- Hofmann, D. J. *Nature* **337**, 447–449 (1989).
- Hofmann, D. J. *Rev. Geophys.* **25**, 743–759 (1987).
- Aimédu, P. & Barat, J. *Rev. Sci. Instr.* **52**, 432 (1981).
- Robbins, D. E. in *Atmospheric Ozone* (eds Zerefos, C. S. & Ghazi, A.) 460–464 (Reidel, Dordrecht, 1985).
- Hofmann, D. J., Harder, J. W., Rolf, S. R. & Rosen, J. M. *Nature* **326**, 59–62 (1987).

- Hofmann, D. J., Harder, J. W., Rosen, J. M., Hereford, J. V. & Carpenter, J. R. *J. geophys. Res.* (in the press).

- Hofmann, D. J., Rosen, J. M. & Harder, J. W. *J. geophys. Res.* **93**, 665–676 (1988).
- Arnold, F. & Knop, G. *Nature* **333**, 746–749 (1989).
- Salawitch, R. J., Gobbi, G. P., Wofsy, S. C. & McElroy, M. B. *Nature* **339**, 525–527 (1989).

ACKNOWLEDGEMENTS. We acknowledge the financial and technical support given to this research by the Centre National d'Etudes Spatiales (CNES). The support of the Centre National de la Recherche Scientifique (CNRS), the NSF Division of Atmospheric Chemistry and Division of Polar programs, NASA, the Chemical Manufacturers Association, the Texas Higher Education Coordinating Board, The Scandinavia-Japan Sasakawa Foundation, Esrange and the Swedish Institute of Space Physics is also gratefully acknowledged. We wish to thank the Stratospheric Group of the Institute for Meteorology of the Free University of Berlin for providing the 30-hPa analyses. D.J.H. acknowledges the support provided by the A. v. Humboldt Foundation and the Institute for Atmospheric Environmental Research, Garmisch-Partenkirchen, FRG while on leave there.

Mapping the transition state and pathway of protein folding by protein engineering

Andreas Matouschek, James T. Kellis Jr, Luis Serrano & Alan R. Fersht*

MRC Unit for Protein Function and Design, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

In the transition state for unfolding of barnase, the hydrophobic core between the major α -helix and β -sheet is somewhat weakened, the C terminus of the major helix is largely intact but its N terminus is exposed and a major loop has been invaded by solvent.

Two fundamental problems to be solved in the mechanism of protein folding are the stability of the folded state and the pathway of folding—in classical terms, the thermodynamics and kinetics. Just as pathways in classical enzyme kinetics have been most rigorously characterized when stable intermediates are present, so progress in understanding protein folding has been most impressive when intermediates have been detected; for example, by trapping transient covalent linkages between cysteines¹ or transient secondary structure by isotope exchange and NMR^{2,3}. Recently, kinetic measurements have been made on the folding of mutant proteins that give information on the roles of specific residues in the processes^{4–7}. We now present a general approach for examining the role of non-covalent interactions in protein folding based on kinetic and thermodynamic methods developed for studying catalysis by the tyrosyl-transferRNA synthetase using protein engineering^{8,9}. The overall strategy is: (1) establish and measure the interactions that stabilize different parts of the molecule from experiments on the equilibrium stability of mutant proteins; (2) use kinetic measurements of protein folding and unfolding to determine what fraction of the stabilization energy measured in 1 is used to stabilize the transition state for folding. This procedure has the potential of discovering the sequence of folding of different parts of the molecule and the interactions in the transition state. Each mutation engineered into the protein acts as a reporter group for the progress of folding at the appropriate position. Crucial to this approach is the choice of mutations. First, as described below, these are strategically located around the molecule to report back on the behaviour of different types of structural elements in the folded enzyme. Second, each mutation was designed to remove a small interaction that stabilizes a discrete region of structure—complex mutations that affect multiple interactions or introduce adverse steric effects were carefully avoided. The effects of more radical and random mutations are often too complicated to interpret. We also show how the changes in kinetics and equilibria of unfolding of mutant enzymes are related and how the relationship may be analysed.

The enzyme studied is barnase, the RNase from *Bacillus amyloliquefaciens* which consists of a single polypeptide chain with no crosslinks. Its virtues for measuring equilibrium interaction energies have been discussed elsewhere^{10–12}. In particular, changes in the free energy of folding on mutation can be accurately measured with little intrusion from artefact¹². It is also a useful system for studying the role of non-covalent interactions in the kinetics of folding. First, there are no disulphide linkages to constrain the unfolded state. Second, the three prolines in barnase, unlike those in RNase A for example, are all *trans*¹³.

The rate-determining step in folding of proteins that contain *cis* prolines in the native structure is sometimes the slow isomerization of proline from the dominant *trans* conformation in the denatured protein¹⁴. Only a small fraction of the denatured barnase exists with *cis* prolines and proline-independent rate constants for unfolding and refolding are readily measured by stopped-flow fluorescence spectroscopy (unpublished data). Here, we are concerned with just measurements of unfolding.

Transition state

The transition state of a single-step chemical reaction is defined simply as the structure at the highest point on the free-energy profile. For a reaction with a series of intermediates, the reaction of each intermediate involves a separate transition state. The accepted convention is that the highest energy transition state is referred to as the transition state for the overall reaction. The transition state for protein unfolding/folding must be an extreme example of a reaction with many intermediates. All physical and spectroscopic evidence on protein structure indicates that individual parts of proteins breathe and reversibly unfold locally at relatively rapid rates. As the structure of the whole of the molecule changes during unfolding and there will be local minima for many of the breathing and local unfolding motions, the highly crenulated free-energy profile for unfolding of Fig. 1 is expected. (This does not contravene the two-state assumption in equilibrium studies because all of the intermediates are at high energy with respect to folded and unfolded states.) The transition state for unfolding is likely to be fuzzy in some parts because the mobile parts of the protein can rapidly interconvert between different conformations of similar energies.

Definitions of rate-determining transition states can become complicated when there are intermediates along the reaction pathway which are more stable than the initial state. This can happen during protein folding^{2,3}. The rate-determining transition state in the reaction is then the highest energy transition

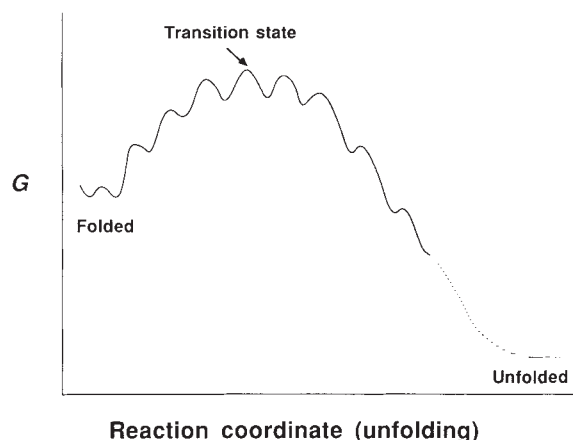


FIG. 1 Illustration of the free-energy profile and transition state for protein unfolding/folding. There are likely to be many small minima during the progress to the highest energy transition state (that is the 'transition state' for the overall reaction) because of the accessibility of different energy states for localized transitions.

* To whom correspondence should be addressed.

state after the stable intermediate. In this case, the kinetics provide information on the energy difference between the transition state and the stable intermediate. To avoid complications in this study from stable intermediates, we have studied the unfolding reaction because we know the structure of the initial and most stable state—that of the native enzyme.

Mutations

Mutations have been prepared that probe various parts of the folded protein (Fig. 2). The hydrophobic core of the protein is formed by the packing of the major α -helix (residues 6–18) on the antiparallel β -sheet. These interactions can be perturbed by changing Leu 14, Ile 88 and Ile 96 to other amino-acid residues by mutation^{10,12}. The protonated imidazole of His 18 interacts with the dipole of the major α -helix at the C cap¹¹ and also makes a hydrogen bond with the backbone NH of Gln 15. Mutation of His 18 is thus a probe of helix formation. The hydroxyl of Thr 6 stabilizes the N terminus of the major helix by acting as a hydrogen-bond acceptor from the backbone >NH groups of Phe 7, Asp 8 and Gly 9 (L.S. and A.R.F.; manuscript in preparation). The hydroxyl of Thr 26 does the same for residues 27, 28 and 29 of the minor α -helix (residues 26–33) (data not shown). New probes constructed here are Tyr $\rightarrow$ Phe 78 that disrupts two hydrogen bonds across a loop of the enzyme to the backbone NH and CO of Gly 81, and Thr $\rightarrow$ Ser 16 on the solvent-exposed surface of the major helix, close to the C terminus of the helix. In all cases, the residues were mutated by making non-disruptive deletions¹⁵ that remove small interactions by making minimal changes in structure that are unlikely to cause serious steric effects. This is even more important in studying the effects of mutation on folding than on catalysis because it is likely that the transition state for folding (Fig. 1) is more easily changed than that for a chemical reaction.

Urea denaturation

Urea denaturation was used in the kinetic and equilibrium measurements. The basic ideas used are those described by

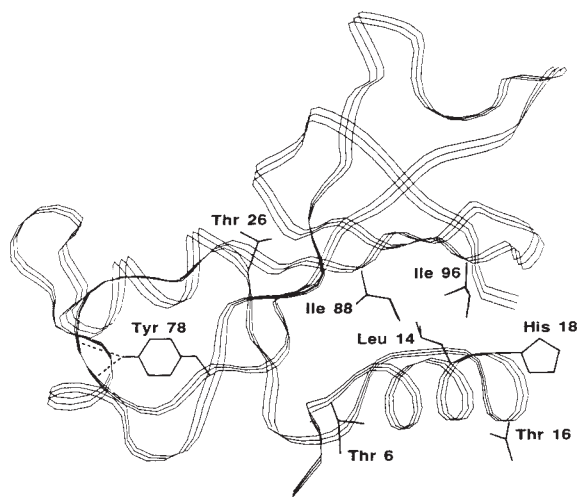


FIG. 2 Sketch of structure of barnase and the residues of interest (produced from coordinates kindly provided by Professor G. Dodson and Dr C. Hill). Thr 6 and Thr 26 act as N caps^{20,21}, accepting hydrogen bonds from the backbone NH groups of principally Gly 9 and Glu 29. The hydroxyl group of Tyr 78 forms hydrogen bonds with the backbone NH and CO groups of Gly 81. The protonated imidazole of His 18 makes an electrostatic interaction with the α -helix dipole, a hydrogen bond with the backbone NH of Gln 15 and hydrophobic interactions with Trp 94 (ref. 11). Ile 88, Ile 96 and Leu 14 form part of the hydrophobic core. Mutant enzymes were prepared as described by Serrano and Fersht (manuscript in preparation) and in refs 10–12. The mutation Tyr $\rightarrow$ Phe 78 was directed by the primer: 5'-ATAATTGAA*ATGTAGTC, and Thr $\rightarrow$ Ser 16 by 5'-TAGAAGTCA*GTATAGTA, where an asterisk follows the mismatch.

Tanford¹⁶. Proteins unfold in urea because side chains and polypeptide backbone are more soluble in urea than in water and unfolding leads to an increase in the exposure of buried structural elements. It is found experimentally that the free energy of transfer of side chains and peptide backbone to urea solutions is approximately linearly proportional to the

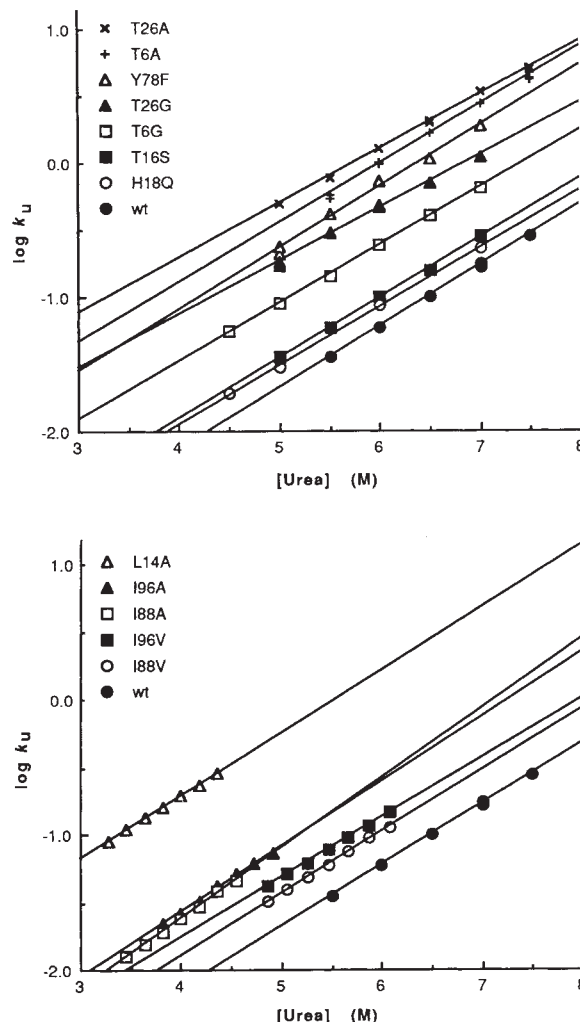


FIG. 3 Plots of $\log k_u$ versus [Urea] for wild-type and mutant barnases. Rate constants measured in units of s^{-1} . All experiments were performed in 50 mM 2-(*N*-morpholino)ethane sulphonic acid buffer (pH 6.3) and at $25 \pm 0.1^\circ C$, monitoring the unfolding by the change in intrinsic fluorescence of the protein ($\lambda_{ex} = 290$ nm, $\lambda_{em} = 315$ nm). Top: unfolding of wild-type enzyme and mutants altered in the helices and loop. Reactions were followed with a Perkin Elmer MPF-44B fluorescence spectrophotometer equipped with a rapid mixing head. The mixing device contained a T-jet mixing chamber followed by a 30-ms delay loop ensuring complete mixing of the solutions before observation. A cylindrical quartz cuvette of 3-mm diameter was used as the observation cell. The solutions were driven through the mixing chamber manually from two Hamilton syringes of different volumes to allow various mixing ratios. Data were acquired with a Tandon Target Microcomputer, a DT2801 Data Translation board and the Bio-logic data acquisition software package by Bio-logic and analysed using the program Enzfitter by R. J. Leatherbarrow. Unfolding was initiated by diluting the aqueous protein solution (~ 0.2 mg ml^{-1}) into urea solutions of different concentrations at a ratio of 1:10. Final concentrations of urea in the observation cell were between 4 and 8 M. Bottom: unfolding of mutants altered in the hydrophobic core. Reactions were followed after manual dilution of barnase (80 μl) into 0.8 ml of buffered urea solution using a Perkin Elmer LS50 fluorescence spectrophotometer (apart from L14A which was monitored as above). Wild-type data from the top are included for reference. The slopes of all lines are within $\pm 5\%$ of 0.45 except for those of Thr $\rightarrow$ Ala 26 and Thr $\rightarrow$ Gly 26 which are 10% lower and Ile $\rightarrow$ Ala 88 which is 16% higher.

concentration of urea ([Urea])¹⁶. This is consistent with the experimental observation that the free energy of unfolding (ΔG_U) is linearly proportional to [Urea]:

$$\Delta G_U = \Delta G_U^{H_2O} - m[Urea] \quad (1)$$

where $\Delta G_U^{H_2O}$ is the free energy of unfolding in water and m is a constant that is proportional to the increase in degree of exposure of the protein on denaturation¹⁶. Rate constants for unfolding and folding also change with [Urea]¹⁶. For example, the rate constant for unfolding, k_u , is often found to increase with increasing [Urea] according to equation 2 where $k_u^{H_2O}$ is the rate constant in H_2O (ref. 16).

$$\log k_u = \log k_u^{H_2O} + m_{k_u}[Urea] \quad (2)$$

In terms of activation energies for the reaction¹⁶:

$$\Delta G_U^\ddagger = (\Delta G_U^{H_2O})^\ddagger - m_u^\ddagger[Urea] \quad (3)$$

The transition state is intermediate between the folded and unfolded states in the degree of exposure to solvent. The fractional increase in exposure of the protein in the transition state for unfolding relative to that on complete unfolding is given by $m_u^\ddagger/m$. This ratio has been used in the past to estimate the average fractional increase in exposure in the transition state¹⁶.

The classical behaviour is followed by the mutants in this study (Fig. 3). In all cases, plots of $\log k_u$ versus [Urea], at concentrations where the enzymes are greater than 95% unfolded at equilibrium, give excellent straight lines which are close to being parallel. The lowered stability of the mutants is reflected in higher rates of unfolding. The average slope (m_{k_u}) is 0.448 which is equivalent to a value of 0.61 for $m_u^\ddagger$. Compared with a value of m for complete unfolding of 2.2, the fractional increase in solvent exposure of the enzyme on going to the transition state is 0.28.

Kinetics and equilibria

The kinetic data show that the mutant enzymes have lower activation energies of unfolding than does wild-type. The change in activation energy is defined by $\Delta\Delta G_U^\ddagger = \Delta G_U^\ddagger - \Delta G_U'^\ddagger$, where $\Delta G_U^\ddagger$ is the activation energy for the unfolding of wild-type enzyme and $\Delta G_U'^\ddagger$ is that for mutant. Experimentally, $\Delta\Delta G_U^\ddagger = -RT \ln(k_u/k_u')$. Important new information about the transition state is available by comparing changes in activation energies from kinetics with overall changes in equilibrium energies. The change in free energy of unfolding is defined by $\Delta\Delta G_U = \Delta G_U - \Delta G_U'$, where ΔG_U is the free energy of unfolding of wild-type enzyme and $\Delta G_U'$ is that of mutant. Experimentally, $\Delta\Delta G_U = -RT \ln(K_U/K_U')$, where K_U is the equilibrium constant for unfolding. It is tempting to take the ratio of $\Delta\Delta G_U^\ddagger/\Delta\Delta G_U$ and equate it directly to β -values in physical-organic chemistry (ref. 17, chapter 2) whereby β is a measure of the extent of bond breaking in the transition state ($\beta = 0$ for no bond breaking, and $\beta = 1$ for complete bond breaking). This cannot be assumed here because here there is an additional term in the equations that is required to account for changes in solvation energy on mutation¹⁸. A full analysis is given in Fig. 4, where $\Delta\Delta G_U^\ddagger$ and $\Delta\Delta G_U$ are related by a pair of thermodynamic cycles, and in equation 4 (compare with ref. 17 chapters 2 and 13):

$$\Delta\Delta G_U^\ddagger/\Delta\Delta G_U = (\Delta G_F - \Delta G_\ddagger)/(\Delta G_F - \Delta G_{solv}) \quad (4)$$

where, as defined in Fig. 4, ΔG_F is the difference in the non-covalent interaction energy between the folded states of wild-type and mutant enzymes, $\Delta G_\ddagger$ between the transition states, and ΔG_{solv} the unfolded states.

ΔG_F , $\Delta G_\ddagger$ and ΔG_{solv} are, of course, global energy terms that refer to the energies of the whole protein. It can be useful, but not essential, for interpretation to attribute ΔG_F and $\Delta G_\ddagger$ mainly to the changes at the bonds disrupted on mutation. This is justifiable when small changes are made that do not cause a

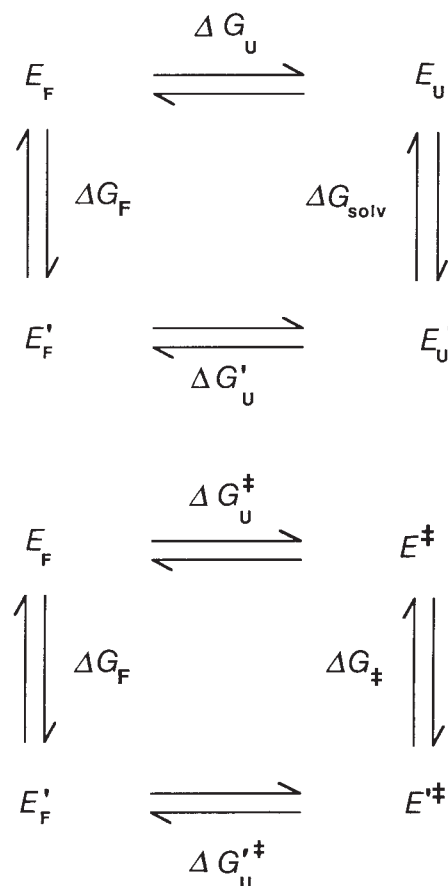


FIG. 4 Thermodynamic cycles that relate free-energy changes in unfolding or wild-type (E) and mutant (E') enzymes. Top, changes in equilibrium unfolding. The difference between the non-covalent energy of folded wild-type and mutant enzymes is defined by $\Delta G_F = G_{E_F} - G_{E'_F}$. Similarly, the difference in the interaction energies between the unfolded proteins and solvents is given by: $\Delta G_{solv} = G_{E_U} - G_{E'_U}$ (compare with ref. 9). The difference in free energy of unfolding $\Delta\Delta G_U = \Delta G_U - \Delta G'_U$, where $\Delta G_U = G_{E_U} - G_{E_F}$. Bottom, the analogous cycle for the formation of the transition states $E^\ddagger$ and $E'^\ddagger$. By transition state theory (see ref. 17); the chemical activation energy, $\Delta G_U^\ddagger = G_{E^\ddagger} - G_{E_F}$. The difference in free energy of activation $\Delta\Delta G_U^\ddagger = \Delta G_U^\ddagger - \Delta G_U'^\ddagger$. The difference between the non-covalent energy of the transition states of wild-type and mutant enzymes is defined by $\Delta G_\ddagger = G_{E^\ddagger} - G_{E'^\ddagger}$. $\Delta G_\ddagger$ and ΔG_F may be related by comparing the two cycles as follows. From the top cycle, $\Delta G_F + \Delta G'_U = \Delta G_{solv} + \Delta G_U$; and from the bottom, $\Delta G_\ddagger + \Delta G_U^\ddagger = \Delta G_F + \Delta G_U^\ddagger$. Therefore, $\Delta\Delta G_U = \Delta G_F - \Delta G_{solv}$, and $\Delta\Delta G_U^\ddagger = \Delta G_F - \Delta G_\ddagger$. The ratio of $\Delta\Delta G_U^\ddagger/\Delta\Delta G_U$ is thus equal to $(\Delta G_F - \Delta G_\ddagger)/(\Delta G_F - \Delta G_{solv})$. Note that the definition of the Brønsted β for unfolding is $\beta = (\Delta G_F - \Delta G_\ddagger)/\Delta G_F$. (β for folding $= \Delta G_\ddagger/\Delta G_F$.) If $\Delta G_{solv} = 0$, then $\Delta\Delta G_U^\ddagger/\Delta\Delta G_U = \beta$ for unfolding.

significant reorganization of the protein¹⁸. Similarly, for a fully denatured protein, ΔG_{solv} is the difference in solvation energy between wild-type side chain and the mutant. The intrusion of ΔG_{solv} in equation 4 is the complicating factor that prevents a simple interpretation. It is possible, however, to find circumstances where $\Delta\Delta G_U^\ddagger/\Delta\Delta G_U$ may be simply interpreted. (To emphasize the difference from β -values, $\Delta\Delta G_U^\ddagger/\Delta\Delta G_U$ is abbreviated to ϕ .) Three examples follow:

(1) *The target region is exposed to solvent in the transition state to the same extent as in the unfolded state.* Under these circumstances, $\Delta G_\ddagger = \Delta G_{solv}$. By inspection of the cycles, $\Delta\Delta G_U = \Delta\Delta G_U^\ddagger$, therefore:

$$\Delta\Delta G_U^\ddagger/\Delta\Delta G_U = \phi = 1 \quad (5)$$

(2) *The interaction energies in the transition states and folded states are very similar.* If $\Delta G_F = \Delta G_\ddagger$, then $\phi = 0$. If $\Delta G_F \approx \Delta G_\ddagger$, and $\Delta G_\ddagger \gg \Delta G_{solv}$, ϕ is small.

TABLE 1 Changes in activation ($\Delta\Delta G_U^\ddagger$) and free energies of unfolding ($\Delta\Delta G_U$) on mutation of barnase

Mutation	Function of position	$\Delta\Delta G_U$ kcal mol ⁻¹	$\Delta\Delta G_U^\ddagger$ (at 4 M urea) kcal mol ⁻¹	$\Delta\Delta G_U^\ddagger$ (in H ₂ O) kcal mol ⁻¹	$\Delta\Delta G_U^\ddagger/\Delta\Delta G_U$ (at 4 M urea)	$\Delta\Delta G_U^\ddagger/\Delta\Delta G_U$ (in H ₂ O)
Thr→Gly 6	N cap*	1.34	0.86	1.01	0.64	0.76
Thr→Ala 6	N cap*	2.23	1.66	1.76	0.74	0.79
Thr→Gly 26	N cap*	1.58	1.33	1.67	0.84	1.06
Thr→Ala 26	N cap*	2.14	1.91	2.19	0.89	1.02
His→Gln 18	C cap—charge/helix dipole	1.60	0.23	0.36	0.14	0.22
Thr→Ser 16	Hydrophobic on helix surface	1.87	0.28	0.37	0.15	0.20
Tyr→Phe 78	Bridges loop	1.50	1.39	1.41	0.92	0.94
Leu→Ala 14	Hydrophobic core	4.80	1.91	1.86	0.40	0.39
Ile→Val 88	Hydrophobic core	1.49	0.30	0.28	0.20	0.19
Ile→Ala 88	Hydrophobic core	4.46	0.68	0.33	0.15	0.07
Ile→Val 96	Hydrophobic core	0.98	0.48	0.55	0.49	0.56
Ile→Ala 96	Hydrophobic core	3.52	0.74	0.60	0.21	0.17

All data at 25 °C and pH 6.3. Values of $\Delta\Delta G_U$ are taken from refs 11, 12 and unpublished data (L.S. and A.R.F.; manuscript in preparation), apart from those for Thr→Ser 16 and Tyr→Phe 78 which were measured as described in ref. 12.* The nomenclature of ref. 21 for residue that occurs at the N terminus of α -helix and can make hydrogen bonds with the backbone NH groups not involved in intrahelical backbone hydrogen bonding²⁰.

(3) $\Delta G_{\text{solv}} = 0$. Mutation of a hydrophilic side chain to a hydrophobic side chain, for example Ser→Ala or Asp→Ala, leads to very large changes in ΔG_{solv} as the hydrophilic groups have large solvation energies, whereas the hydrocarbon groups do not. Conversely, $\Delta G_{\text{solv}} \approx 0$ for small changes in aliphatic hydrophobics (for example, Ile→Ala or Val have values of ΔG_{solv} in water for fully exposed side chains of only -0.21 and -0.16 kcal per mol respectively)¹⁹. Then for mutations such as Ile→Ala or Val and the side in chains in the unfolded structure are exposed to water, ΔG_{solv} can be ignored when ΔG_F is significant and:

$$\Delta\Delta G_U^\ddagger/\Delta\Delta G_U = \phi = (\Delta G_F - \Delta G_T)/\Delta G_F \quad (6)$$

Relationship between β and ϕ . The above examples are simple to interpret because β and ϕ are, in fact, identical for those circumstances. In case 3, $\Delta G_{\text{solv}} = 0$, which is the condition for $\phi = \beta$ under all conditions (see legend to Fig. 4). In case 1 where there is complete bond breaking in the transition state, $\beta = 1$ by definition, and ϕ is seen to be equal to 1 (equation 5). In case 2 where there is no bond breaking in the transition state, $\beta = 0$ by definition, and also $\phi = 0$. But, in cases other than 1 or 2 where $\Delta G_{\text{solv}} \neq 0$, $\phi \neq \beta$. There will be an increasing discrepancy between ϕ and β in cases 1 and 2 as ϕ differs further from 0 and 1, and will be particularly large where hydrophilic groups are mutated to hydrophobic. Another discrepancy where $\phi \neq \beta$ arises when there is residual structure in the unfolded protein because ΔG_{solv} will then contain the energetics of those structural interactions. Barnase was chosen as there is minimal likelihood of residual structure in its unfolded states, and this assumption is supported by the finding that there are identical free-energy changes of folding measured by thermal-, urea- and guanidinium hydrochloride-induced denaturation¹². Given this reasoning and assumptions, we can now describe the transition state for unfolding at the positions where mutations have been made.

Unfolding

The ratios $\Delta\Delta G_U^\ddagger/\Delta\Delta G_U$ (that is, ϕ) are listed in Table 1. There are only minor differences between the ratios observed at 4 M urea and those extrapolated to 0 M urea, indicating that there are no artefacts from the binding of urea to specific groups of the protein. The interpretations divide into the three classes above:

(1) *Altered region is exposed to solvent in the transition state*, $\Delta\Delta G_U^\ddagger/\Delta\Delta G_U \approx 1$. The side chains of the threonines in the N caps form hydrogen bonds with the backbone NH groups at the beginning of the helices. The values of ϕ are close to unity for the mutation Thr 26→Ala or Gly, and the values for Thr 6 mutations are very high. Either the helices are unwound in the transition state in these regions or just the N termini are exposed. (As seen below, it is likely that the helices still exist).

Alteration of Tyr→Phe 78 also has $\phi \approx 1$. In global energy terms this means that the portion of the loop is as exposed to solvent in the transition state as is the unfolded state. In specific terms, the hydrogen bonds between the tyrosyl hydroxyl and Gly 81 that stabilize the folded structure are broken in the transition state and hydrogen bonds are made with water.

It could be argued that the ϕ -values of 1 arise not because the loop has become invaded with solvent but because hydrogen bonds in the loop or N cap are stretched in the transition state and weakened by an energy which is, fortuitously, the same as that observed for complete unfolding. If this is so then there are two states available to the loop which are of equal energy—strained or exposed to solvent. Given the free-energy profile of Fig. 1, it would be expected that both states would be equally occupied in the transition state as they are of the same energy. Thus, a ϕ -value of 1 in this context implies that at least 50% of the loop is exposed to solvent in the transition state.

(2) *Transition state has largely the folded structure*, $\phi \approx 0$. His 18 interacts with the electrostatic dipole at the C terminus of the major α -helix (residues 6–18). The ϕ -value for mutation of His→Gln 18 is only ~ 0.2 . This indicates that the interaction is largely intact in the transition state. For this to occur, the helix must still be present in the transition state. Further evidence for the integrity of the C terminus of the helix is given below.

(3) *Hydrophobic interactions*, $\Delta G_{\text{solv}} \approx 0$. The methyl group of Thr 16 makes hydrophobic interactions with the aromatic ring of Tyr 17 in the C terminus of the helix which are lost on unfolding. The ϕ -value of 0.15–0.2 means, according to equation 6, that 80–85% of the non-covalent interaction energy is maintained in the transition state. This is additional evidence that the helix is largely intact in the transition state.

Changing the hydrophobic residues in the core gives ϕ -values in the range 0.3 ± 0.2 . Thus some 70% of the interaction energy remains in the transition state. The core is weakened, but still has significant energy.

In conclusion, the transition state for unfolding has significant secondary and compact structure. The major α -helix (residues 6–18) is largely intact, the hydrophobic core is weakened but still important, and the major α -helix is still docked on the central β -sheet. One loop is considerably exposed to solvent in the transition state as are the N termini of both helices.

Folding pathway

The transition state for folding is, by the principle of microscopic reversibility, the same as that for unfolding under identical conditions and for a reversible process. We have thus shown that the structure of the transition state for folding is like the structure of the folded state, with much of the secondary structure having been formed and with significant interactions in the hydrophobic core. These results are consistent with recent NMR

evidence that during folding an early framework intermediate forms in which there is significant secondary structure^{2,3} and the proposal that the transition state occurs late on the folding pathway¹. The initial results from our protein engineering study have begun to map out the energetics of these processes and to show new details such as a major loop not being formed in the

transition state or the lack of hydrogen bonds from N caps to the exposed backbone NH groups of the helices. These studies now point to new questions that can be answered by studying further mutations. For example, are all the loops exposed to solvent in the transition state and are they the last elements of structure to be formed? □

Received 20 April; accepted 7 June 1989.

1. Creighton, T. E. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5082-5086 (1988).
2. Udgaonkar, J. B. & Baldwin, R. L. *Nature* **335**, 694-699 (1988).
3. Roder, H., Elöve, G. A. & Englander, S. W. *Nature* **335**, 700-704 (1988).
4. Beasty, A. M. *et al. Biochemistry* **25**, 2965-2974 (1989).
5. Garvey, E. P. & Matthews, C. R. *Biochemistry* **28**, 2083-2093 (1989).
6. Chen, B.-I., Baase, W. A. & Schellman, J. A. *Biochemistry* **28**, 691-699 (1989).
7. Goldenberg, D. P., Frieden, R. W., Haack, J. A. & Morrison, T. B. *Nature* **338**, 127-132 (1989).
8. Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M. & Smith, M. *Nature* **299**, 756-758 (1982).
9. Fersht, A. R. *Biochemistry* **26**, 8031-8037 (1987).
10. Kellis, J. T., Jr, Nyberg, K., Sali, D. & Fersht, A. R. *Nature* **333**, 784-786 (1988).
11. Sali, D., Bycroft, M. & Fersht, A. R. *Nature* **335**, 741-743 (1988).
12. Kellis, J. T., Jr, Nyberg, K. & Fersht, A. R. *Biochemistry* **28**, 4914-4922 (1989).

13. Mauguen, Y. *et al. Nature* **297**, 162-164 (1982).
14. Schmid, F. X. & Baldwin, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4764-4768 (1978).
15. Fersht, A. R., Leatherbarrow, R. J. & Wells, T. N. C. *Biochemistry* **26**, 6030-6038 (1987).
16. Tanford, C. *Adv. Protein Chem.* **24**, 1-95 (1970).
17. Fersht, A. R. *Enzyme Structure and Mechanism* 2nd edn (Freeman, New York, 1985).
18. Fersht, A. R. *Biochemistry* **27**, 1577-1580 (1988).
19. Wolfenden, R., Andersson, L., Cullis, P. M. & Southgate, C. C. B. *Biochemistry* **20**, 849-855 (1981).
20. Presta, L. G. & Rose, G. D. *Science* **240**, 1632-1641 (1988).
21. Richardson, J. S. & Richardson, D. C. *Science* **240**, 1648-1652 (1988).

ACKNOWLEDGEMENTS. J.T.K. is a recipient of a NATO postdoctoral fellowship. L.S. is a recipient of an EMBO postdoctoral fellowship.

LETTERS TO NATURE

Nucleosynthesis, neutrino bursts and γ -rays from coalescing neutron stars

David Eichler*, Mario Livio†, Tsvi Piran‡
& David N. Schramm§

* Department of Physics, Ben Gurion University, Beer Sheva, Israel, and Astronomy Program, University of Maryland, College Park, Maryland 20742, USA

† Department of Physics, The Technion, Haifa, Israel

‡ Racah Institute for Physics, Hebrew University, Jerusalem, Israel, and Princeton University Observatory, Princeton, New Jersey 08544, USA

§ Departments of Physics and Astrophysics, University of Chicago, 5640 Ellis Avenue, Chicago, Illinois 60637, USA, and NASA/Fermilab Astrophysics Center, Batavia, Illinois 60510, USA

NEUTRON-STAR collisions occur inevitably when binary neutron stars spiral into each other as a result of damping of gravitational radiation. Such collisions will produce a characteristic burst of gravitational radiation, which may be the most promising source of a detectable signal for proposed gravity-wave detectors¹. Such signals are sufficiently unique and robust for them to have been proposed as a means of determining the Hubble constant². However, the rate of these neutron-star collisions is highly uncertain³. Here we note that such events should also synthesize neutron-rich heavy elements, thought to be formed by rapid neutron capture (the r-process)⁴. Furthermore, these collisions should produce neutrino bursts⁵ and resultant bursts of γ -rays; the latter should comprise a subclass of observable γ -ray bursts. We argue that observed r-process abundances and γ -ray-burst rates predict rates for these collisions that are both significant and consistent with other estimates.

The binary pulsar system will coalesce in roughly 10^8 yr. Using this fact together with the pulsar birth rate and the observed frequency of occurrence of close binary pulsars (then, one in every 300 pulsars), Clark *et al.*³ estimated the formation rate in the Galaxy to be $(3 \pm 1.6) \times 10^{-4} \text{ yr}^{-1}$. As 450 pulsars have now been observed, this estimate should be decreased by a factor of 1.5, but this does not change the following discussion. The natural progenitor for a close neutron-star binary system is a massive X-ray binary, from which the neutron-star binary will form following a supernova explosion. Clark *et al.*³ estimated that the probability of disruption of such a close binary in the

supernova of the secondary is 0.1-0.2, and this led them to a formation rate consistent with that based on pulsar statistics.

Alternatively, neutron stars formed in separate events might become associated in a dense globular cluster^{6,7}. Such modes of formation will lead to roughly equal-mass binaries with both neutron stars of mass $1.4 M_{\odot}$.

On the other hand, were there to exist a separate class of neutron-star binaries in which the stars are in close proximity on formation of the binary, these might escape detection because they would decay on a short timescale, τ_{gr} , via gravitational radiation. For example, it has been suggested recently⁶ that millisecond pulsars arise from accretion-induced collapse of white dwarfs in binaries⁷. It is argued on statistical grounds⁶ that low-mass X-ray binaries are too rare to be the progenitors of millisecond pulsars, as had been proposed previously⁸. As the white dwarfs are likely to be spun up to maximum angular momentum, the collapse to a neutron star is likely to 'fizzle', leading to two neutron stars in orbit around each other⁹. It is not unlikely that the two neutron stars would have significantly different masses. Although the statistical arguments have been criticized (F. Verbunt, J. van Paradijs and W. Lewin, personal communication), such criticism does not exclude the possibility that close neutron-star binaries are indeed formed from white dwarfs in globular clusters.

Essentially for all equations of state, the less massive component (the secondary) has a larger radius¹⁰, R_2 . As it also has the smaller Roche-lobe radius, R_L , it is invariably this lighter component that fills its tidal lobe first, following which mass transfer ensues. If $\zeta_2 \equiv (\partial \ln R_2 / \partial \ln M_2) > \zeta_L \equiv (\partial \ln R_L / \partial \ln M_2)$ the mass transfer is dynamically stable, whereas for $\zeta_2 < \zeta_L$ it is unstable. The latter situation arises when the secondary ('donor') neutron star has a mass that is comparable to or only slightly smaller than that of the larger, whereas a high initial mass ratio will result in a stable mass-stripping process.

Clark and Eardley¹ have discussed the evolution of a close neutron-star binary. In systems that are unstable to mass transfer on a dynamical timescale, the margin by which the less massive neutron star is the first to overfill its lobe grows until the orbital evolution is dominated by mass exchange (rather than by gravitational radiation), for which the timescale¹¹ Δt_{ex} is ~ 6 ms for masses of $M_1 = 1.4 M_{\odot}$, $M_2 = 1.2 M_{\odot}$. Once the mass-transfer timescale becomes shorter than the gravitational-radiation timescale, mass loss accelerates rapidly. In a three-dimensional simulation of this process for a doubly degenerate binary system, it is found that the lighter component is completely dissipated in a little more than two orbital periods (~ 4 ms in our case).

The secondary is transformed into a thick, axially symmetric disk orbiting around the primary. Such configurations for the case of a white-dwarf binary have been constructed recently¹². In the calculation of Benz *et al.*¹³, about 0.3% of the total mass escaped the system. A similar fraction can be expected in our case, because this fraction depends mainly on the surface potentials of the two stars and is thus proportional to $(1 - M_2 R_1 / M_1 R_2)$.

Calculations^{12,13} have shown that, for the white-dwarf binary, the collapse does not occur immediately even though the total mass exceeds the Chandrasekhar mass, mainly because the system is supported by centrifugal forces (stability is also indicated by the ratio of kinetic to gravitational energy). Thus, the late-stage evolution of the configuration depends critically on transport and disposal of angular momentum from the disk. If transport of angular momentum occurs entirely by means of degenerate matter viscosity¹⁴, then extremely long viscous timescales result. The value of the Reynolds number indicates, however, that turbulence may arise in a significant fraction of the disk¹². In this case, the relevant viscosity may be the turbulent viscosity

$$\nu_{\text{turb}} \approx (2 \times 10^9 \text{ cm}^2 \text{ s}^{-1}) (V_t / 10^8 \text{ cm s}^{-1}) (l_t / 10^5 \text{ cm}) (\text{Re}_c / 5,000)^{-1},$$

where l_t and V_t are the typical size and velocity of a turbulent cell and Re_c is the critical Reynolds number. Such a turbulent disk will transport angular momentum on a timescale

$$\tau_{\text{vis}} \approx 500 (R / 10^6 \text{ cm})^2 (\nu_{\text{turb}} / 2 \times 10^9 \text{ cm}^2 \text{ s}^{-1})^{-1} \text{ s}.$$

If the binary is stable ($\zeta_2 > \zeta_1$) against dynamical-timescale mass transfer (that is, if the initial mass ratio is large), the mass-transfer rate, $\dot{M}_2$, will approach asymptotically a value dictated essentially by gravitational radiation¹⁵: $\dot{M}_2 \approx -2M_2 / (\zeta_2 - \zeta_1) \tau_{\text{gr}}$. This accretion rate is enormous and it raises again the question of how angular momentum is disposed of and of whether the accreting neutron star has the ability to accommodate such a large mass influx. In the case of stable mass transfer there is another interesting possibility. Neutron stars have a minimum mass below which they are unstable to free expansion. In the present context, the secondary may be stripped to this minimum mass before colliding with the more massive component, at which stage it will explode¹⁶⁻¹⁸.

An interesting feature of the decompression of the secondary is the accompanying de-neutronization of the material. One expects the neutrons to aggregate into large droplets, in the same way as they do in decompressed water^{19,20}. Assuming that the difference in Fermi energies of the neutron and proton in the neutron-rich nuclei is of the order of a few MeV, the β -decay timescale is of the order of a few milliseconds near the neutron drip line (where neutron-rich nuclei are no longer able to capture or retain additional neutrons), which is comparable to the decompression timescale. Hence the de-neutronization is non-adiabatic. This has two important implications. First, decompression can occur while the nuclei are far from the β -decay instability, and thus the nuclei may undergo rapid neutron capture^{4,21}. It has been pointed out^{19,20} that such a process can produce r-process nuclei in the observed mass range with abundance peaks corresponding to the neutron 'magic' numbers²². This is particularly interesting because these abundance peaks are observed to be quite narrow and to be distinct from the corresponding s-process abundance peaks. More conventional r-process models involving supernova explosions require fine tuning of temperature and neutron densities throughout many different dynamical events in order to produce such sharp peaks^{23,24}. We see, however, that neutron-star decompression may give rise to sharp peaks as a consequence of the intrinsic nuclear physics rather than having to rely on hydrodynamic concordance throughout different events.

Neutron capture is extremely rapid and throughout most of the decompression the nuclei remain very close to the neutron drip line. However, because the β -decay timescale is then

comparable to the expansion timescale, radioactivity causes significant heating, so that the temperature may be of the order of 1 MeV throughout most of the decompression phase and the actual neutron excess may be somewhat below that of the zero-temperature neutron drip line, as required by the location of the r-process peaks.

Furthermore, antineutrinos should be emitted at energies of up to a few MeV. The β -decay states are usually excited states with energies again of a few MeV, and rapidly emit γ -ray cascades as they decay to stable states.

In a close neutron-star binary, the larger neutron star accretes most of its companion over a dynamical timescale of $\sim 10^{-2}$ s, liberating energy at a rate of $\sim 10^{54}$ erg s⁻¹, and copious neutrino emission is expected. The rate of mass infall onto the primary is within an order of magnitude of that expected from spherically symmetric core collapse, so the neutrino spectrum which is buffered by energy-dependent interaction cross-sections, should resemble the (now measured) burst from a core collapse. Note that in the coalescing case the matter in the envelope is hotter than in the isolated core collapse and it is therefore easier for the neutrinos to diffuse. The neutrino diffusion timescale is of the order of 1 s, but because the distribution of infalling material is biased towards the orbital plane, neutrinos may escape out of the plane on a significantly shorter timescale. A total energy output of $\sim 10^{53}$ erg is expected from the accretion process, a sixth of which is due to electron neutrinos (unfortunately, such neutrinos cannot be detected with existing detectors beyond ~ 100 kpc). An energy of $\sim 10^{51}$ erg is expected from antineutrinos that result from the de-neutronization of the companion during its decompression.

The electron neutrinos from the accreting primary and the antineutrinos from the disrupted secondary can produce electron-positron pairs where they collide out of the orbital plane. The efficiency with which antineutrinos are converted is roughly²⁵

$$\varepsilon \approx 7 \times 10^{-3} \alpha f [(\sigma_{\nu\bar{\nu}} / 10^{-44} \text{ cm}^2) (\bar{E}_\nu / 10 \text{ MeV})^2] \times \left[\frac{(E_\nu / 10^{53} \text{ erg})}{(\bar{E}_\nu / 10 \text{ MeV})} \right] (R / 10^6 \text{ cm})^{-1} (\tau / 0.1 \text{ s})^{-1}$$

where $\sigma_{\nu\bar{\nu}}$ is the cross-section for $\nu\bar{\nu} \rightarrow e^+e^-$, E_ν is the total energy emitted in neutrinos, $\bar{E}_\nu$ is the average energy of a neutrino; R is the size of the region in which the neutrino flux is large and τ is the duration of the neutrino burst; α is a geometrical factor that depends on the geometry of the neutrinosphere and f is the filling factor of the system.

The total energy converted to e^+e^- pairs and, ultimately, the time profiles of the neutrino emission make this estimate of ε rather uncertain. The energy of the primary γ -ray cascades accompanying the β -decay antineutrinos should be within an order of magnitude of this total energy, that is, $\geq 10^{50}$ erg.

A comparable or larger γ -ray flux might be produced from pair production by $\nu\bar{\nu}$ pairs emerging from, say, opposite sides of the primary because, although the geometry is less favourable, the greater total energy and the greater compactness may more than compensate. Here the key uncertainty concerns the matter that presumably envelopes the more massive neutron star as it accretes onto it. Because this matter is intensely heated by its own radioactivity, we cannot draw too many parallels with previously catastrophic-mass-transfer models. In order that the γ -rays emerge before complete degradation, they must be produced beyond the accretion column, and this may require too small an intersection angle between the neutrino and antineutrino trajectories to sustain a centre-of-mass energy much above $2 m_e c^2$.

The γ -ray fireball has such a large compactness parameter that any soft photons produced within are Compton-upscattered and kinetic equilibrium is attained with the remaining pairs. The fireball is optically thick for temperatures $T > 10^9$ K, and e^+e^- pairs are produced. The fireball will expand until most of

the pairs have annihilated irreversibly, and it then becomes optically thin. Although the photons are redshifted during the expansion in the frame of the co-moving fluid, the bulk expansion at the onset of transparency causes a blueshift in the frame of the observer by the same factor. The overall spectrum will be a modified black-body spectrum with the same temperature as initially but with fewer hard photons²⁶. A large uncertainty in the end result of this process stems from the potential appearance of a thick baryonic wind. If a large amount of baryonic matter is mixed in with the fireball, most of the available energy will be used in accelerating it and the γ -ray signal will then not be so strong.

For a luminosity of 10^{47} erg s⁻¹ emitted from a region of size 2×10^7 cm, the temperature of the region is $\sim 10^9$ K. The spectrum and timescale should be similar to those of a typical γ -ray burst. As the Universe is transparent to soft γ -rays^{27,28}, it should be detectable to distances of ~ 100 Mpc with fluences typical of γ -ray bursts (10^4 eV cm⁻²), if the energy output is $\sim 10^{46}$ erg. However, we do not expect such radiation to have cyclotron absorption features, nor a significant non-thermal tail above several MeV. These two features are relatively common for γ -ray bursts, so that the scenario we propose can explain only a subset of them, that is, 'featureless' γ -ray bursts. Such featureless bursts would be extragalactic and thus isotropic. Statistics on featureless γ -ray bursts may thus provide an estimate of the rates of neutron-star coalescence events.

In numerical models of cataclysmic mass transfer, a small fraction ($\sim 0.3\%$) of the transferred mass is expelled to infinite distances. It is possible that radioactive heating, which can supply an energy of $\sim 3 \times 10^{-3} Mc^2$, can cause most of the disrupted star to explode and escape to infinity. There is $\sim 10^5 M_\odot$ per galaxy of material produced by the r-process, and a few times $10^3 M_\odot$ of the heavy r-process material that might be favoured in this type of event. If we assume that $\sim 0.3\%$ of the mass is ejected per event, then such events should occur at a rate of 3×10^5 to 3×10^7 times per galaxy per Hubble time (the lower limit applies if we wish to explain the heavy r-process material in this way and the upper limit applies if we wish to explain all r-processed material). These rates bracket the rate obtained by Clark *et al.*³ of 3×10^{-4} events per year per galaxy. If a subset of these events is also to be able to account for millisecond pulsars, then we require only about 10^{-5} events per year, or 10^5 per galaxy per Hubble time. Here we have assumed an average spin-down time for a typical millisecond pulsar of 10^9 years, and that there are currently about 10^4 millisecond pulsars in our Galaxy.

This scenario makes two simple observational predictions. First, assuming that $\sim 10^5$ galaxies are within 100 Mpc and that the bursts are indeed detectable out to that distance, then an occurrence of $\sim 10^{-4}$ per galaxy per year yields a detection rate of 10 per year. With the oriented scintillation spectrometer experiment on the Gamma Ray Observatory, it will be relatively straightforward to distinguish featureless, highly thermal γ -ray bursts from others. Should such a class be identified, we suggest that it would be worthwhile to check for identifications of such bursts with galaxies. Second, gravitational-radiation events of this nature should be detectable with a 30σ signal up to a distance of 100 Mpc and with a 3σ signal up to a distance of 1,000 Mpc by the proposed Caltech-MIT Gravitational Wave Detector². The rate of stronger events should be comparable to that of γ -ray bursts of this kind, and the coincidence of such γ -ray bursts with gravity waves may in fact provide the most stringent observational test of the scenario. Verification would imply that our model identifies the site of the astrophysical r-process. $\square$

- Clark, J. P. A., van den Heuvel, E. P. J. & Sutantyo, W. *Astr. Astrophys.* **72**, 120-128 (1979).
- Lattimer, J. & Schramm, D. N. *Astrophys. J.* **210**, 549-556 (1976).
- Berezinsky, V. & Prilutski, P. *Proc. 19th Int. Cosmic Ray Conf.* **1**, 29-32 (La Jolla, 1985).
- Grindlay, J. E. & Bailyn, C. D. *Nature* **336**, 48-50 (1988).
- Ostriker, J. P. *Sources of Gravitational Radiation* (ed. Smarr, L. L.) 461-476 (Cambridge University Press, 1979).
- Alpar, M. A., Cheng, A. F., Ruderman, M. A. & Shaham, J. *Nature* **300**, 728-730 (1982).
- Clark, J. P. A. *Sources of Gravitational Radiation* (ed. Smarr, L. L.) 447-459 (Cambridge University Press, 1979).
- Arnett, W. D. & Bowers, R. L. *Astrophys. J. Suppl.* **33**, 415-436 (1977).
- Webbink, R. F. *Astrophys. J.* **277**, 355-360 (1984).
- Mochkovitch, R. & Livio, M. *Astr. Astrophys.* **209**, 111-118 (1989).
- Benz, W., Bowers, R. L., Cameron, A. G. W. & Press, W. H. *Astrophys. J.* (in the press).
- Itoh, N., Kohyama, Y. & Takenchi, H. *Astrophys. J.* **317**, 733-736 (1987).
- Webbink, R. F. & Iben, I. Jr in *IAU Colloq. 95, The Second Conference on Faint Blue Star* (eds Davis Philip, A. G., Hayes, D. S. & Liebert, J. W. 778 (L. Davis Press, Schenectady, 1987).
- Page, D. N. *Phys. Lett. A* **91**, 201-202 (1982).
- Shapiro, S. in *Numerical Astrophysics* (eds Centrella, J., Leblanc, J. & Bowers, R.) 190, 215 (Bertlett, Boston, 1985).
- Colpi, M., Shapiro, S. L. & Teukolsky, S. A. *Astrophys. J.* (in the press).
- Lattimer, J. M., Mackie, F., Ravenhall, D. G. & Schramm, D. N. *Astrophys. J.* **213**, 225-233 (1977).
- Meyer, B. *Astrophys. J.* (in the press).
- Symbolist, E. M. D. & Schramm, D. N. *Astrophys. Lett.* **22**, 132-145 (1982).
- Burbidge, E. M., Burbidge, C. A., Fowler, W. A. & Hoyle, F. *Rev. mod. Phys.* **29**, 547-650 (1957).
- Norman, E. & Schramm, D. N. *Astrophys. J.* **228**, 881-892 (1979).
- Cameron, A. G. W., Cowan, J. & Truran, J. in *Nucleosynthesis* (eds Arnett, W. D. & Truran, J.) 190-201 (University of Chicago Press, 1985).
- Berezinsky, V. & Prilutski, F. *Proc. 20th Int. Cosmic Ray Conf.* **1**, 54-68 (Moscow, 1987).
- Goodman, J. *Astrophys. J.* **308**, L47-L50 (1986).
- Babul, A., Paczynski, B. & Spergel, D. N. *Astrophys. J.* **316**, L49-L54 (1987).
- Shapiro, S. L. & Teukolsky, S. A. *Black Holes, White Dwarfs and Neutron Stars* (Wiley, New York, 1983).

ACKNOWLEDGEMENTS. We thank J. Goodman, B. Paczynski and B. Meyer for discussions. This work was supported in part by NASA and the NSF at Chicago, by the NASA Fermilab Astrophysics Center at Maryland and by BSF at Jerusalem. We thank the Jerusalem Winter School, the IAS at Princeton (D.N.S. and D.F.E.) and the Princeton University Observatory (T.P.) for hospitality.

Observation of interfacial atomic steps during silicon oxidation

J. M. Gibson & M. Y. Lanzerotti*

AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974, USA

SILICON oxidation is perhaps the most common process in the fabrication of electronic devices. The silicon/silicon dioxide interface has excellent electrical properties which are employed in field-effect devices and surface passivation, and are arguably responsible for the dominance of silicon technology in the microelectronics industry. As device dimensions shrink, the need for high-quality oxides grown at increasingly lower thicknesses and temperatures has demanded ever more perfect, extremely flat Si/SiO₂ interfaces. Here we use transmission electron microscopy to image exposed and buried monatomic interfacial steps at the Si/SiO₂ interface during oxidation. The results show that extremely flat boundaries can be formed at low temperatures and that atomic steps do not play a dominant role in the oxidation process itself. A simple bulk-terminated Si/amorphous SiO₂ interface is consistent with our data.

The nature of the terminating Si lattice at the Si/SiO₂ boundary influences the transport of carriers in metal-oxide-semiconductor field-effect transistors (MOSFETs), and correlation has been observed between apparent interface roughness and Hall mobility¹. In the high-resolution transmission electron microscopy (HRTEM) study of ref. 1 it was emphasized that only statistical properties of interface roughness could be identified because of the implicit projection through ~ 100 Å specimen thickness. One detailed study² of Si/SiO₂ interfacial roughness with low-energy electron diffraction (LEED) assumed that the Si/SiO₂ interfacial structure was unperturbed by stripping of the oxide in HF. Other experiments³⁻⁷ which examined the stoichiometry of the SiO_x layer adjacent to the interface suggest an abruptly terminating Si lattice. In a recent HRTEM study of Si layers and their native oxides⁸ deposited by molecular beam

Received 20 March; accepted 20 May 1989.

- Clark, J. P. A. & Eardley, D. M. *Astrophys. J.* **215**, 311-322 (1977).
- Shutz, B. F. in *Gravitational Collapse and Gravity* (eds Sato, H. & Nakamura, T.) 246-253 (Singapore World Scientific, 1986).

* Present address: Dunster House, Harvard University, Cambridge, Massachusetts 02138, USA.

epitaxy (MBE), it was claimed that a crystalline intermediate phase of SiO_2 (tridymite) exists at the Si/SiO₂ interface formed by this technique.

Here we present images and diffraction patterns from clean Si surfaces before, during and after the formation of a native oxide layer, obtained by ultra-high-vacuum TEM. We use a technique that involves surface-sensitive Bragg reflections to study buried interfaces, and are thus able to monitor terrace and step distributions at the buried Si/SiO₂ interface.

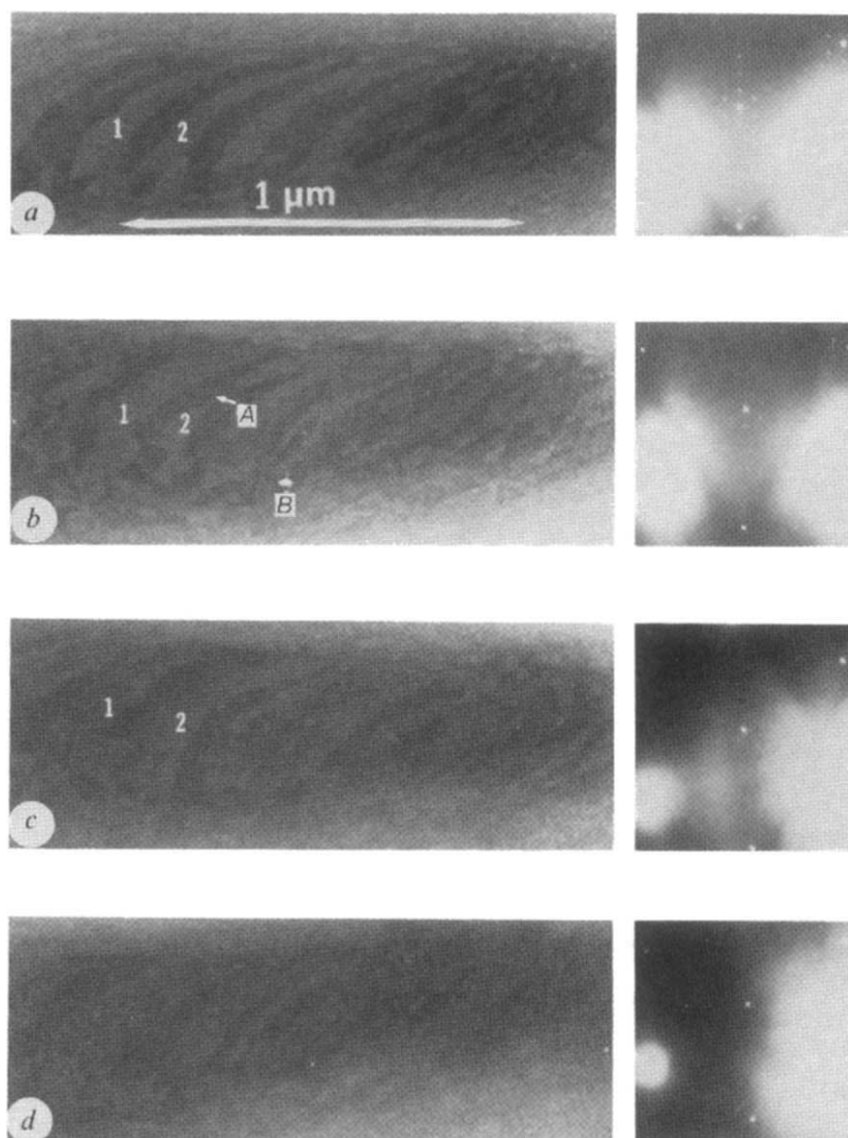
Bragg reflections of type $1/3\ 422$ in (111) TEM diffraction patterns were first observed from thin gold foils⁹. These reflections are forbidden for an infinite lattice, but they may persist when a non-integral number of unit cells in a thin specimen causes incomplete cancellation of the diffraction from hexagonally symmetric (111) layers, or equivalently from a reciprocal lattice spike associated with a 111 reflection in the first-order Laue zone⁹. Kinematical diffraction theory predicts an oscillation in the intensity of this reflection (as a result of the large effective deviation parameter $s_z = 1/(\sqrt{3}a_0)$) with thickness period $\sqrt{3}a_0$, so that images of gold foils with monolayer contours of foil thickness could be observed¹⁰. Cherns has pointed out that in (100) foils, similar effects should occur at the 110 diffraction position, and this was later observed in gold¹⁰. Recently, analogous diffraction effects have been studied in the scattering of X-rays from surfaces^{11,12}.

Iijima¹³ observed sharp $1/3\ 422$ spots after heat treatment of (111) thin specimens of silicon at 1,200 °C in a separate ultra-high-vacuum system. In general, the $1/3\ 422$ spots are not seen from 'typical' Si surfaces prepared by chemical polishing, although they have been observed from 'exceptionally thin and flat' specimens¹⁴. When stacking faults are present, these can greatly enhance the intensity in the $1/3\ 422$ reflection^{14,15}. In these previous studies the air-exposed Si samples had native oxides on the surface. Our experiments indicate that $1/3\ 422$ or 110 reflections from Si surfaces prepared by ion-thinning or HF/HNO₃ etching are invisible in typical 100-kV transmission electron diffraction patterns, whereas they are seen from clean surfaces. This can be explained by the effect of surface roughness on the intensity distribution of these forbidden reflections. For a gaussian, randomly stepped surface it can be shown that the diffraction profile near the $1/3\ 422$ position is¹⁶

$$I(s_x) = \frac{\alpha/L}{\pi^2/(9L^2) + s_x^2}$$

where s_x is the component of the deviation parameter perpendicular to [111] and L is the expected spacing between steps of monolayer height ($a_0/\sqrt{3}$). The peak half-width increases and intensity drops as L decreases, so for heavily stepped surfaces the spot can become invisible against the background. Experimentally this occurs for step spacings of less than 20 Å. Only

FIG. 1 Images taken with $1/3\ 422$ reflections, revealing atomic terraces on a thin ($\sim 1,000$ -Å thickness) Si specimen in an ultrahigh-vacuum TEM at 100 kV. Diffraction patterns are inset, showing *a*, the clean Si (111) 7×7 superstructure (with the $1/3\ 422$ diffraction spot at the centre); *b*, partial transformation to a 1×1 surface after exposure to 600 monolayers O₂; *c*, the 1×1 unreconstructed nature of the native oxide/Si interfaces; and *d*, the effect of stripping the oxide in an aqueous HF solution. The images reveal the behaviour of atomic steps. The significance of arrows and numerals is discussed in the text.



after high-temperature oxidation and annealing, or *in situ* heat treatment, are the steps far enough apart for clear diffraction to be evident (Fig. 1). The spot profile provides a microscopic method of quantifying roughness at buried interfaces¹⁶.

Perforated silicon specimens were cleaned *in situ* by heating to $\sim 1,200^\circ\text{C}$ in an ultra-high-vacuum TEM operated at 100-kV with a base pressure of 1×10^{-9} torr. Oxygen was introduced through a controlled leak valve. Figure 1 shows $1/3$ 422 dark-field images and inset diffraction patterns from the same area of a (111) Si sample during exposure to O_2 . Figure 1a reveals the step distribution on the clean 7×7 reconstructed surface at room temperature. The boundaries between lighter (for example, regions 1 and 2 in Fig. 1) and darker terrace regions represent steps on either of the two specimen surfaces. Two-thirds of all steps are observed—those that change thickness from $3n+1$ to $3n+2$ layers are invisible. The average step separation on a single surface is approximately $2,000 \text{ \AA}$. Before oxygen exposure, the sample was cleaned by flash heating to $1,000^\circ\text{C}$, which caused very little redistribution of steps. Figure 1b was taken after *in situ* exposure to 5×10^{-6} torr O_2 for 300 s (~ 600

monolayers). The 7×7 diffraction pattern has weakened, indicating a surface of mixed 7×7 and 1×1 regions. The step distribution, although still recognizable from Fig. 1a, has changed noticeably. Some steps have moved by $\sim 500 \text{ \AA}$ (at arrow A, for example, two steps have come together). Figure 1c was recorded from the same specimen area after exposure to ambient air at atmospheric pressure for 15 h, creating a native oxide of thickness $\sim 20 \text{ \AA}$ (ref. 17). Steps and terraces at the buried Si/SiO₂ interfaces are visible and, surprisingly, are still recognizable from the distribution in Fig. 1a and b. Finally, Fig. 1d shows the result of stripping the sample in a 5% HF/H₂O solution for 5 min, which is expected to completely remove the oxide, with no effect on the silicon substrate¹⁸. The interface is still atomically flat but the step distribution is very different from the buried configuration.

Only two intensity levels occur in most terrace images of Fig. 1, implying that the surface and Si/SiO₂ interface always terminate at the same Si (111) double-layer¹⁴. In Fig. 1b, small bright spots are observed (for example, arrow B), of approximate dimension $50 \text{ \AA}$. We suggest that these are imperfectly stacked islands of Si debris, created by the interaction of oxygen with the 7×7 reconstructed surface and similar to the Si/As surface¹⁹. These features disappear after formation of native oxide.

Figure 2 shows diffraction patterns from a (001) surface before and after exposure to O_2 . The *in situ* cleaned surface shows type 110 spots (for example, 110, $\bar{1}30$), as expected from crystal termination effects, and type $1/2, 1/2, 0$ spots (triangular arrowheads) from the 2×1 surface reconstruction. In general, the 110 spots on clean surfaces are streaked, indicating a higher step-density than (111) surfaces after similar high-temperature treatment (as a result of streaking, these spots are difficult to see in the high background, but intensity measurements are reported below). Figure 2b was taken after exposure of the same area to air at atmospheric pressure and room temperature for 12 h. The 110 spots are still visible. Step spacings of $\sim 50 \text{ \AA}$ are found on our high-temperature cleaned (001) Si surfaces, just at the practical resolution limit for step imaging. Sharper 110 diffraction spots have been seen after high-temperature oxidation of (001) samples¹⁶, indicating that the Si/SiO₂ interface can be flatter than the clean surface.

The measured $1/3$ 422 intensities in the (111) diffraction patterns of Fig. 1a–c are a, $(5.4 \pm 1) \times 10^{-4}$, b, $(6.0 \pm 2) \times 10^{-4}$ and c, $(2.5 \pm 1) \times 10^{-4}$. (Absolute values are accurate only to within a factor of two. Errors refer to relative values.) We do not find any great sensitivity of these intensities to crystal orientation (away from strong Bragg excitations), as has been reported by others with less-flat specimens^{13,14}. From an ideally terminated Si(111) surface we calculate an average kinematical $1/3$ 422 intensity at 100 kV of $\sim 1 \times 10^{-3}$, which could be doubled by dynamical diffraction¹³. The reduced intensity observed on oxidation is consistent with the expected 'shuffle' termination of the Si/SiO₂(111) interface in which only one bond per atom is made with oxygen⁶. On the (001) surface of Fig. 2 the 110 (or 310) intensity is $(1.5 \pm 0.4) \times 10^{-4}$ for the clean surface (a) and $(7.0 \pm 3) \times 10^{-5}$ after exposure to atmospheric O_2 and consequent native-oxide formation (b). The calculated value for a perfectly terminated surface is 1.2×10^{-4} . Thus there is no evidence of the increased intensity that would result from epitaxial phases at either the $1/3$ 422 (on (111)) or 110 (on (100)) positions after native-oxide formation. Note that five monolayers of well-ordered tridymite, which have been suggested to be present in similar circumstances⁸, should cause an increased intensity of 1.9 – 3.5×10^{-4} from each surface (A. Ourmazd, personal communication). However, it is likely that the MBE-grown Si surface⁵ is flatter than our high-temperature-cleaned surface. It has been suggested²⁰ that crystalline phases develop only in oxides thicker than $20 \text{ \AA}$. We have also observed that conventional high-temperature-grown oxides with very flat interfaces exhibit termination Bragg reflections with no evidence of additional intensity from ordered epitaxial oxides¹⁶.

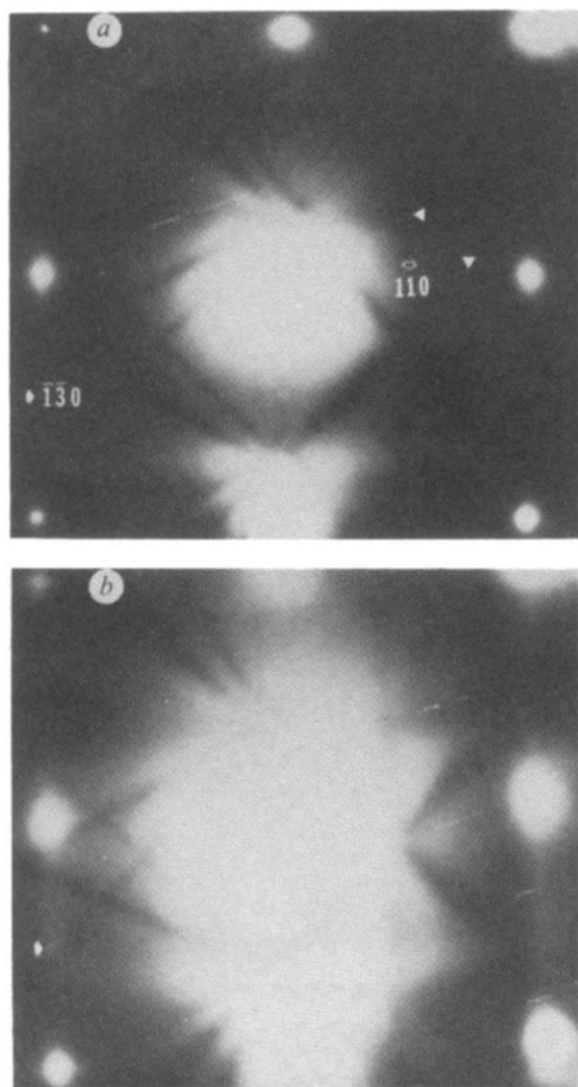


FIG. 2 100-kV transmission electron diffraction patterns from: a, clean (001) 2×1 reconstructed surfaces of a thin Si specimen, and b, the surface after native-oxide formation. The 2×1 superstructure has disappeared but the 110 type spots remain at the buried interfaces, indicating the presence of flat and perfectly terminated Si terraces. Triangular arrowheads are discussed in the text.

Although some steps have moved considerable distances, the observation that step patterns are recognizable after native oxidation of a flat Si surface implies that the growth of the ~ 20 -Å-thick native oxide does not occur by a terrace/ledge mechanism. For a single monolayer of growth by this mechanism, a step must propagate across the entire surface, and this should randomize the step distribution. To explain the removal of stress, Mott²¹ has proposed that oxidation of Si occurs preferentially at steps and kink sites, and Mazur and Washburn²² have also inferred from HRTEM studies that ledges and kinks are involved in oxidation. Our results suggest that this mechanism is not dominant, and that instead oxygen must attack Si randomly on the terraces with no strong preference for steps at the buried interface, in agreement with a study of Si etching by oxygen²³.

After attack on the oxide layer by a 5% HF/H₂O solution, it has been found that the Si surface not only is stripped of oxide but remains so for a considerable period of time in air¹⁸. Hahn and Henzler's² LEED studies of interface roughness assumed

that the (pure) HF-stripped surface was structurally equivalent to the buried interface. Our data validate this assumption by showing that there is substantial step redistribution on stripping the oxide in aqueous HF, but that the step density is not dramatically affected.

In conclusion, we show step arrangements on clean Si (111) surfaces and their behaviour on formation of the native oxide, which implies that the oxidation process is not dominated by a terrace/ledge-kink mechanism at room temperature. Extremely flat Si/SiO₂ interfaces result from our room-temperature oxidation experiments on clean, flat Si surfaces. The use of forbidden reflections in producing these images provides a wealth of information about the structure of buried interfaces, as originally proposed by Cherns⁹. The appearance and sharpness of forbidden reflections is a very sensitive function of surface/interface roughness. Our *in situ* data also show that atomically flat Si terraces can exist at SiO₂/Si (111) and (100) interfaces with no intermediate epitaxial phases present. □

Received 21 April; accepted 2 June 1989.

- Goodnick, S. M. *et al. Phys. Rev. B* **32**, 8171–8185 (1985).
- Hahn, P. O. & Henzler, M. *J. Vac. Sci. Technol. A* **2**, 574–583 (1984).
- Aspnes, D. E. & Theeten, J. B. *J. electric. Soc.* **127**, 1359–1365 (1980).
- Jackman, T. E., McDonald, J. R., Feldman, L. C., Silverman, P. J. & Stensgaard, I. *Surf. Sci.* **100**, 35–42 (1980).
- Hollinger, G. & Himpfel, F. *J. Appl. Phys. Lett.* **44**, 93–95 (1983).
- Grunthaner, P. J., Hecht, M. H., Grunthaner, F. J. & Johnson, N. M. *J. appl. Phys.* **61**, 629–638 (1987).
- Grovenor, C. R. M. *Mat. Res. Soc. Proc.* **53**, 301–310 (1986).
- Ourmazd, A., Taylor, D. W., Rentschler, J. A. & Bevk, J. *Phys. Rev. Lett.* **59**, 213–216 (1987).
- Cherns, D. *Phil. Mag.* **30**, 549–556 (1974).
- Krakow, W. & Ast, D. G. *Surf. Sci.* **58**, 485–496 (1976).
- Andrews, S. R. & Cowley, R. A. *J. Phys. C* **18**, 6427–6439 (1985).
- Robinson, I. K., Waskiewicz, W. K., Tung, R. T. & Bohr, J. *Phys. Rev. Lett.* **57**, 2714–2717 (1986).
- Iijima, S. *Ultramicroscopy* **6**, 41–52 (1981).

- Ourmazd, A., Anstis, G. R. & Hirsch, P. B. *Phil. Mag. A* **48**, 139–153 (1983).
- Alexander, H., Spence, J. C. H., Shindo, D., Gottschalk, H. & Long, N. *Phil. Mag. A* **53**, 627–643 (1986).
- Gibson, J. M., Elser, V. & Lanzerotti, M. Y. *Appl. Phys. Lett.* (submitted).
- Carim, A. H., Dovek, M. M., Quate, C. F., Sinclair, R. & Vorst, C. *Science* **237**, 630–633 (1987).
- Grunder, M. & Schulz, R. *J. Vac. Sci. Technol.* (in the press).
- Becker, R. S., Swartzentruber, B. S., Vickers, J. S., Hybertson, M. S. & Louie, S. G. *Phys. Rev. Lett.* **60**, 116–119 (1988).
- Ross, F. M. & Stobbs, W. M. *Mat. Res. Soc. Proc.* **105**, 259–264 (1988).
- Mott, N. F. *Proc. R. Soc. A* **376**, 207–215 (1981).
- Mazur, J. H. & Washburn, J. *Am. Inst. Phys. Conf. Ser.* **122**, 52–55 (1984).
- Shimizu, N., Tanishiro, Y., Kobayashi, K., Takayanagi, K. & Yagi, K. *Ultramicroscopy* **18**, 453–462 (1985).

ACKNOWLEDGEMENTS. We are grateful for assistance from M. L. McDonald and D. Bahnck and discussions with V. Elser, A. Howie and G. S. Higashi.

Non-steady-state biological removal of atmospheric particles from Mediterranean surface waters

P. Buat-Ménard*, J. Davies*, E. Remoudaki†, J. C. Miquel‡, G. Bergametti†, C. E. Lambert*, U. Ezat*, C. Quétel*, J. La Rosa‡ & S. W. Fowler‡

* Centre des Faibles Radioactivités, Laboratoire Mixte CNRS-CEA, Domaine du CNRS, F-91198 Gif sur Yvette Cedex, France

† Laboratoire de Physico-Chimie de l'Atmosphère, Université Paris 7, 2 Place Jussieu, F-75251 Paris Cedex 05, France

‡ International Atomic Energy Agency, International Laboratory of Marine Radioactivity, 2 Avenue Prince Héritaire Albert, MC-98000, Monaco

It has recently been suggested that the deposition of mineral aerosol particles, primarily transported from Africa during sporadic but intense events^{1,2}, may have a profound influence on the geochemistry and the sedimentology of the Mediterranean Sea^{1–3}. Several studies in the open ocean^{4–7} have shown that such particles, with radii generally less than 5 µm, are removed from the surface waters within a few weeks, together with large organic particles produced by biological packaging⁷. Here we report data from time-series measurements of both atmospheric inputs and water-column particulate fluxes at 200 m depth obtained from sediment traps, which show that such a biological control also prevails in northwestern Mediterranean waters. On short time-scales, concentrations and fluxes in the upper water column can increase significantly following dust transport and deposition events, with a response time of the order of one week. Such non-steady-state behaviour must be taken into account when assessing the impact of pulsed atmospheric inputs on particulate trace element concentrations and fluxes in the water column.

Sampling and analysis of total (wet + dry) atmospheric deposition samples for aluminium, an indicator of aluminosilicate minerals, have been described elsewhere¹. Samples have been collected with a hemispheric plastic collector (of area 0.1 m²) at Capo Cavallo (42° 31' N, 8° 40' E), Corsica, since February 1985. This site is located 300 m above sea level and is ~ 700 m distant from the shore line. The sampling duration was ~ 15 days. As part of the Dynamique des Flux Atmosphériques en Méditerranée programme, mineral fluxes in the water column have been obtained since April 1986 from an automated time-series cylindrical particle-trap^{8,9} (0.125-m² opening and height-diameter aspect ratio of 2.5; Technipac PPS-3) moored at a depth of 200 m in a 2,200-m water column, 15 nautical miles off the coast of Calvi, Corsica (42° 44' N, 8° 31' E); that is, ~ 20 nautical miles from the atmospheric sampling site. Each of the six trap collector cups contained a buffered formalin solution and was set to sample sinking particles for consecutive periods of 6–15 days. After retrieval of the traps, the particulate material was immediately processed on return to shore. A small aliquot ($\frac{1}{8}$ – $\frac{1}{6}$ of the total sample) was sampled for direct optical observations of the collected particulate material. The remaining fraction was freeze-dried and weighed. Seawater samples were collected in Go-Flo bottles during several cruises to obtain vertical profiles of dissolved and particulate materials. After collection, samples were passed through pre-weighed Nucleopore filters (0.4-µm pore size). The aluminium concentration in the filters and the trap material was determined by instrumental neutron-activation analysis¹⁰. Scanning electron microscopy and energy-dispersive X-ray spectrometry shows that most of the aluminium in the filter or trap samples is associated with aluminosilicate materials, in agreement with previous studies^{6,11}.

Figure 1 shows the temporal variation of the aluminium atmospheric deposition flux for 1985–87 and we see that the deposition fluxes are highly variable, ranging from 0.14 to 18.6 mg m⁻² d⁻¹. The highest deposition fluxes correspond to the so-called 'red rain' episodes in this region. Three-dimensional air-mass trajectories, as well as satellite imagery,

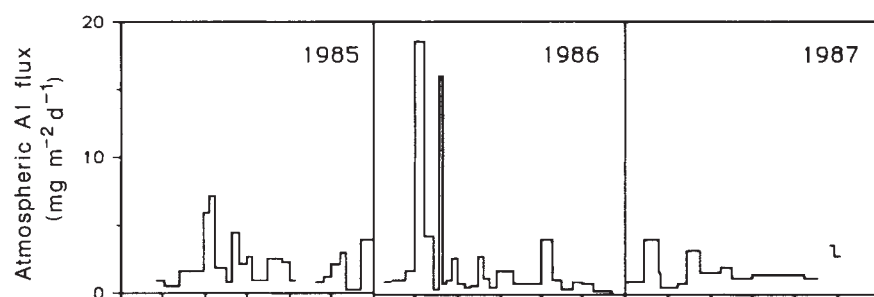


FIG. 1 Temporal variations of aluminium atmospheric deposition at Capo Cavallo, Corsica. The interval marked on the x axis corresponds to a two-month period. Mean deposition fluxes for 1985, 1986 and 1987 are 1.84, 2.05 and 1.56 $\text{mg m}^{-2} \text{d}^{-1}$ respectively.

indicate that such episodes are associated with the transport of soil dust from Africa¹. For example, 20 dust events, each lasting about 2 days, were recorded between April 1985 and April 1986. These events contribute most of the annual atmospheric deposition of mineral aerosol particles in this marine region, with one in March 1986, accounting for 30% of the annual deposition flux. Such variability in dust transport and deposition clearly requires long time-series measurements to assess the geochemical significance of atmospheric inputs. In spite of this, mean aluminium deposition fluxes showed little interannual variability from 1985 to 1987 (1.84, 2.05 and 1.56 $\text{mg m}^{-2} \text{d}^{-1}$ respectively). The magnitude of the deposition flux in this region agrees with the estimate of Loÿe-Pilot *et al.*² based on rain samples collected in southern Corsica.

Table 1 gives the results from the sediment-trap samples. They represent the particulate fluxes through 200 m during a 5-week sampling period in spring 1986 and during 205 days in 1987. The sampling depth of 200 m below the surface was chosen as it has been suggested that vertical fluxes of organic matter passing that depth should be equivalent to new production^{12,13}. The range of mass fluxes that we observed, 8–335 $\text{mg m}^{-2} \text{d}^{-1}$, is within the range of results obtained for low-productivity open ocean waters¹³. The highest mass fluxes are observed during later winter and spring when primary productivity generally

increases in the northwestern Mediterranean¹⁴. Based on a mean crustal Al abundance of 8.2% (ref. 15), we calculated that aluminosilicate minerals accounted for between 5 and 76% of the total mass flux whereas, based on Ca determinations by neutron-activation analysis, we estimated that calcium carbonate represented between 12 and 65% of the flux. Thus, a major fraction of the dry mass flux, between 40 and 90%, was due to the presence of these two types of material. Organic carbon was only measured on trap samples collected during sampling periods 1, 2 and 3 and represented between 8 and 15% of the dry mass flux.

The particulate aluminium flux at 200 m ranged between 0.03 and 11.7 $\text{mg m}^{-2} \text{d}^{-1}$. This range is much smaller than that found by other studies in the Mediterranean Sea^{9,16} in which advective inputs from the coast were significant. The mean value observed, 2.1 $\text{mg m}^{-2} \text{d}^{-1}$, compares favourably with the mean atmospheric deposition flux of 1.8 $\text{mg m}^{-2} \text{d}^{-1}$, suggesting that on a yearly basis, atmospheric deposition at our sampling site can support the water-column flux of mineral particles. This hypothesis is supported by the consideration of the relative abundance of aluminium in the individual trap samples. Indeed, the highest aluminium abundances were generally found shortly after a period of high atmospheric deposition (Fig. 2).

Ten vertical profiles of suspended particulate aluminium were obtained between April 1986 and October 1987. In the upper 200 m, the standing crop of particulate aluminium (average of 4–10 samples) varied between 35 and 130 mg m^{-2} (J.D. *et al.*, manuscript in preparation). Consideration of the mean standing crop in the upper 200 m (80 mg m^{-2}) suggests a mean residence time of 40 days for mineral particles with respect to both mean atmospheric input and sediment-trap flux data which agree with results from studies in the North Atlantic^{5,6}. The mean residence time corresponds to an apparent mean sinking rate of $\sim 5 \text{ m d}^{-1}$, which far exceeds those of individual mineral particles (with radii generally $< 5 \mu\text{m}$) based on stokesian settling calculations. Numerous studies^{7,11} have shown that aluminosilicate minerals in sea water are efficiently scavenged by organic aggregates and/or incorporated into a faecal pellets which have sinking velocities of the order of tens of several hundreds m d^{-1} . Microscopic examination of the trap samples during sampling periods 1 and 2 indicate that they were rich in faecal material typical of that from copepods. This material made up between 10 and 70% of the total dry weight in these samples (S.W. and J.C.M., manuscript in preparation). For these two sampling periods, we observed a significant correlation ($r = 0.91$) between the particulate aluminium flux and the faecal pellet flux (Fig. 3). This degree of correlation is much higher than that observed between particulate aluminium and total mass or organic carbon fluxes, which would suggest that zooplankton grazing and subsequent faecal pellet production may be the most efficient removal process of mineral particles from these surface waters. This was shown to be the case during spring 1986 for Chernobyl fallout radionuclides⁸, in particular the rare-earth nuclides Ce-141 and Ce-144, which entered the Mediterranean as a single pulse and were rapidly removed from surface waters and transported to 200 m in a few days. Similar conclusions were reached to explain the temporal variability of the natural radionuclide Th-234 at our trap site¹⁷. Such rapid removal probably explains the

TABLE 1 Results from particle interceptor trap experiments

Sample	Period	Mass flux ($\text{mg m}^{-2} \text{d}^{-1}$)	Al flux ($\text{mg m}^{-2} \text{d}^{-1}$)	% Al
1986				
1-1	13–20 Apr.	214	3.5	1.6
1-2	20–26 Apr.	112	2.4	2.2
1-3	26 Apr.–2 May	64	1.7	2.8
1-4	2–8 May	66	ND	ND
1-5	8–15 May	54	1.0	1.8
1-6	15–21 May	57	0.88	1.3
1987				
2-2	18–27 Jan.	71	0.51	0.72
2-3	27 Jan.–5 Feb.	151	2.7	1.8
2-4	5–14 Feb.	144	8.9	6.2
2-5	14–23 Feb.	335	11.7	3.5
2-6	23 Feb.–4 Mar.	276	2.4	0.87
3-1	27 Mar.–6 Apr.	49	0.71	1.5
3-2	6–16 Apr.	172	3.0	1.7
3-3	16–26 Apr.	160	3.2	2.0
3-4	26 Apr.–6 May	126	1.9	1.5
3-5	6–16 May	217	4.1	1.9
4-1	22 June–2 July	236	3.4	1.5
4-2	2–12 July	134	1.1	0.78
4-3	12–22 July	76	0.43	0.57
4-4	22 July–1 Aug.	55	0.57	1.0
4-5	1–11 Aug.	41	0.63	1.5
4-6	11–21 Aug.	13	0.10	0.79
5-1	11–21 Sep.	30	0.31	1.0
5-2	21 Sep.–1 Oct.	8.2	0.031	0.38
5-3	1–11 Oct.	12	0.2	1.65
5-4	11–21 Oct.	51	1.2	2.4
5-5	21 Oct.–1 Nov.	45	1.2	2.7

ND, not determined.

observed increases in particulate aluminium concentrations in the individual trap samples following periods of high atmospheric deposition.

Such a complex interaction can be illustrated by the data obtained during sediment-trap sampling period 2 from 18 January to 4 March 1987 (Fig. 2). At the beginning of this period, the aluminium atmospheric deposition rate was low ($<1 \text{ mg m}^{-2} \text{ d}^{-1}$). It then increased by a factor of 4 between 27 January and 17 February and decreased again afterwards. Also a sharp increase in the aluminium trap flux is noted between 5 and 23 February (Table 1, Fig. 2) which suggests an influence of the atmospheric pulse on the aluminium trap flux. However, for this whole sampling period the aluminium flux recorded in the traps (230 mg m^{-2}) was found to be about twice as large as the aluminium atmospheric deposition rate (105 mg m^{-2}). As this period was also characterized by a significant increase in the faecal pellet flux (Fig. 2), this result can be explained by the fact that the zooplankton biomass graze not only atmospheric particles delivered by the recent atmospheric pulse but also 'older' atmospheric particles which, because of previously lower biological activity, had not been removed from surface waters. Within the limitation of our data set, the suspended aluminium standing crop of 67 mg m^{-2} observed on 3 February 1987, before the atmospheric pulse, is sufficient to provide the additional aluminium required to support our hypothesis.

The mean standing crop of particulate aluminium on 27 February 1987 was 86 mg m^{-2} thus only slightly higher than on 3 February. This implies a very fast removal of the atmospheric pulse, less than two weeks. On shorter timescales we should nevertheless expect a large increase of the suspended aluminium

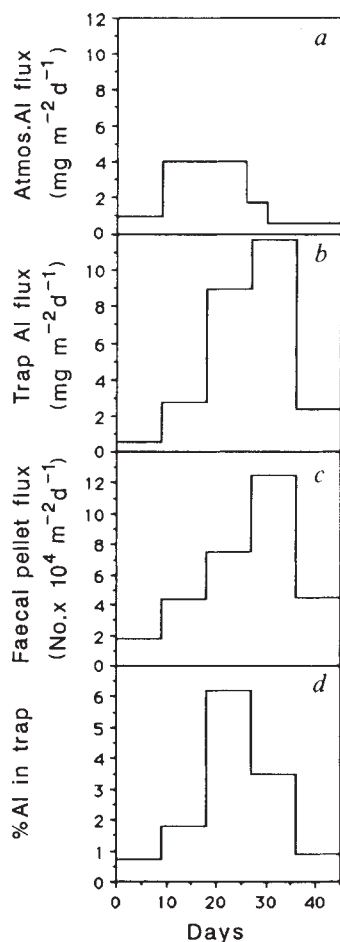


FIG. 2 Temporal variations of a, aluminium atmospheric deposition; b, aluminium sediment-trap flux; c, faecal pellet fluxes; and d, aluminium relative abundance (%) in trap samples from 18 January to 4 March 1987.

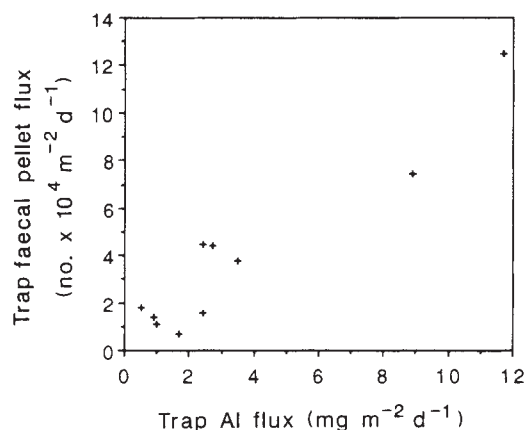


FIG. 3 Aluminium versus faecal pellet fluxes during sediment-trap sampling periods 1 and 2 off the coast of Corsica. The regression line is $y = 0.661 + 0.927x$, where y is the faecal pellet flux and x the aluminium flux.

standing crop, by a factor of two to four, for several days following an atmospheric pulse of the same intensity. Such data are not available for this sampling period. We have, however, observed changes of this magnitude or greater within a few days in spring 1988 at the same site, following an African dust transport episode (J.D. *et al.*, manuscript in preparation). Such short-term changes have also been observed in North Atlantic and North Pacific surface waters^{18,19}.

There is therefore a dominant biological, and presumably seasonal, control of the removal of atmospheric particles from these surface waters. The fact that yearly atmospheric inputs are balanced by removal with the large particle flux clearly suggests that direct transfer of atmospheric particles to deep waters by convective mixing was not important during the time interval considered here. Our results also suggest a variable mean response time of the deep water column to inputs from the atmosphere. Indeed, depending on the removal rate, the mean residence time of particulate aluminium in the upper 200 m can vary by almost an order of magnitude, from a value of ~ 10 days to 100 days or more. This implies that particulate aluminium is never at steady state on timescales of a month or less. This non-steady-state behaviour with respect to removal has evidently to be taken into account when assessing the impact of pulsed inputs from the atmosphere on particulate aluminium concentrations and fluxes in the water column. Non-steady-state models must therefore be developed for the prediction of the fate of atmospheric inputs of particulate trace elements in surface and deep waters. For a given marine region, this will require an adequate parametrization of the atmospheric input function and the use of realistic pelagic ecosystem models²⁰. Such a modelling approach is currently being applied (ref. 21 and D. Ruiz-Pino *et al.*, manuscript in preparation). □

Received 20 March; accepted 31 May 1989.

- Bergametti, G. *et al.* in *Paleoclimatology and Paleometeorology: Modern and Past Patterns of Global Atmospheric Transport* (eds Leinen, M. & Sarinheim, M.) (Reidel, Dordrecht, in the press).
- Loje-Pilot, M. D., Martin, J. M. & Morelli, J. *Nature* **321**, 427–428 (1986).
- Buat-Ménard, P. in *The Role of Air-Sea Exchange in Geochemical Cycling* (ed. Buat-Ménard, P.) 477–496 (Reidel, Dordrecht, 1986).
- Honjo, S. *Science* **218**, 883–884 (1982).
- Wallace, G. T., Mahoney, O. M., Dulmage, R., Storti, F. & Dudek, N., *Nature* **293**, 729–731 (1981).
- Deuser, W. G., Brewer, P. G., Jickells, T. D. & Commeau, R. F. *Science* **219**, 388–391 (1983).
- Fowler, S. W. & Knauer, G. A. *Progr. Oceanogr.* **16**, 147–194 (1986).
- Fowler, S. W. *et al.* *Nature* **329**, 56–58 (1987).
- Monaco, A. *et al.* in *Particle Flux in the Oceans* (ed. Degens, E. T., Izdar, E. & Honjo S.) 109–125 (SCOPE/UNEP, 1987).
- Delmas, R., Dulac, F., Lambert, C. E. & Revel, G. *J. radioanalyst. Nucl. Chem.* **111**, 371–383 (1987).
- Lambert, C. E., Jehanno, C., Silverberg, J. S., Brun-Cottan, J. C. & Chesseelet, R. *J. mar. Res.* **39**, 77–98 (1981).
- Eppley, R. W. & Peterson, B. J. *Nature* **282**, 677–680 (1979).
- Hargrave, B. T. *Ecol. Model.* **30**, 229–246 (1986).
- Jacques, G. *Oceanol. Acta Spec. Iss.* No. 9, 143–148 (1988).
- Taylor, S. R. *Geochim. cosmochim. Acta* **28**, 1273–1285 (1964).
- Copin-Montégut, C., *Oceanol. Acta Spec. Iss.* No. 9, 95–102 (1988).

17. Buat-Ménard, P. et al. in *Radionuclides, a Tool for Oceanography* (ed. Guay, J. C. Guegueniat, P. & Pentreath, R. J.) 121–130 (Elsevier, London, 1988).
18. Uematsu, M., Duce, R. A., Nakaya, S. & Tsunogai, S. *J. geophys. Res.* **90**, 1167–1172 (1985).
19. Chester, R. in *The Role of Air-Sea Exchange in Geochemical Cycling* (ed. Buat-Ménard, P.) 443–476 (Reidel, Dordrecht, 1986).
20. Andersen, V. & Nival, P. *Mar. Ecol. Prog. Ser.* **44**, 37–50 (1988).
21. Ruiz-Pino, D., Lambert, C. E., Jeandel, C. & Buat-Ménard, P. *Chem. Geol.* **70**, 198 (1988).

ACKNOWLEDGEMENTS. This research was supported in part by CNRS, CEA, University of Paris 7 and the French Institut des Sciences de L'Univers (Programme Flux Océaniques, opération DYFAMED), the Programme Interdisciplinaire de Recherches sur l'Environnement (PIREN-CNRS, ATP 'Aérosols Désertiques'), the French Ministry of the Environment, the World Meteorological Organisation, the (VERTEX program) and the UNEP MEDPOL Program. We thank A. Vigot, P. Bourdon, J. Farber-Lorda, the crew and officers of RV *Korotneff*, RV *Catherine Laurence* and RV *Recteur Dubuisson*, the French National Navy for use of the sampling facilities at Capo Cavallo and the Laboratoire Pierre Sûe at Saclay for assistance and use of facilities in activation analysis. The International Laboratory of Marine Radioactivity operates under an agreement between the International Atomic Energy Agency and the principality of Monaco.

Grain-boundary graphite in rocks and implications for high electrical conductivity in the lower crust

B. Ronald Frost*, William S. Fyfe†, Kazue Tazaki† & Tammy Chan†

* Department of Geology and Geophysics, University of Wyoming, Laramie, Wyoming 82071, USA

† Department of Geology, University of Western Ontario, London, Ontario N6A 5B7, Canada

THE origin of zones of high electrical conductivity in the lower continental crust is a long-standing mystery; possible explanations include the presence of brines^{1,2}, partial melt³, serpentine⁴ and graphite⁵. When discussing the occurrence of graphite in the crust many petrologists have considered phase relations as they would have existed at the peak of metamorphism or during igneous emplacement^{6–10}. Here we show that a fine film of graphite is present on the grain boundaries in three rocks from the Laramie Anorthosite Complex. In two of these rocks graphite was stable during igneous crystallization but in the other it was not. We maintain that in all of the rocks the grain-boundary graphite precipitated from a CO₂-rich fluid during cooling. The chemical processes that produced the grain-boundary graphite in these rocks are likely to operate in many lower-crustal rocks. We therefore contend that, because the films we observe are capable of producing the high conductivity that is seen in the lower crust, grain-boundary graphite should be considered as a possible cause for some conductivity anomalies.

To test whether graphite can form a continuous grain-boundary film in rocks, we studied samples from the 1.4-Gyr-old Laramie Anorthosite Complex. Samples from three rock types were studied; two (W-34-5-2, LAC-8) came from the Sybille Monzosyenite, and the other (LAC-10B) came from a particularly fresh sample of gabbroic anorthosite. The Sybille Monzosyenite was chosen because it is known to have crystallized in the stability field of graphite¹¹. The rock is a late-stage pluton in the Laramie Anorthosite Complex, consisting of alkali feldspar, fayalite, hedenbergite, quartz and ilmenite. Phase equilibria indicate that the rock crystallized at pressures of around 3 kbar and at temperatures between 950 and 1030 °C (ref. 11). Oxygen fugacity was near that of graphite saturation, between 1 and 2 log units below the FMQ buffer (see Fig. 7 of ref. 11). Fluid-inclusion and Raman studies indicate that crystallization took place in the presence of graphite and a nearly pure CO₂–CO vapour phase¹². The gabbroic anorthosite contains cumulus plagioclase with a post-cumulus assemblage of olivine (Fa₇₀)–augite–orthopyroxene ($X_{Fe} = 0.58$, $X_{Fe} = Fe/(Fe + Mg + Mn)$; inverted from pigeonite)–ilmenite–magnetite, with ilmenite far more abundant than magnetite. This assemblage probably formed at 1,100 °C and at oxygen fugacities around 1 log unit below FMQ.

Scanning Auger multiprobe analysis was carried out with a Perkin-Elmer Physical Electronics model 600 instrument with an accelerating voltage of 2.5–3.0 keV. Gently crushed bulk samples were mounted on indium foil and were analysed after a 30-s sputtering to remove atmospheric contamination of carbon. Auger-spectrometer depth profiling showed a high concentration of carbon on grain boundaries (the sensitivity of detection is around the 1% level) (Fig. 1). Scanning Auger carbon-content maps were made with an accelerating voltage of 3 keV and beam diameter of about 1 µm without any coating. Sputter rates of ~600 Å min⁻¹ for SiO₂ and a 4-kV argon-ion beam were used. The carbon map clearly shows high carbon



FIG. 1 a and b, Scanning electron microscopy photographs for a sample of the Sybille Monzosyenite (sample LAC-8). The numeral denotes the locality for the graphite spectrum given in Fig. 2. c, Auger spectrometry carbon map for the same sample. Scale is the same as in b.

intensity at the grain boundaries as substantiated by the point analysis. The carbon spectrum from this analysis (Fig. 2) is identical to pure crystalline graphite from Sri Lanka and is quite distinct from carbonate carbon and from contamination carbon from the vacuum system. Sputtering indicates that the graphite films are $\sim 1,000$ Å thick.

To determine whether graphite films of this thickness are sufficient to produce the high conductivity seen in the lower crust, the expression in ref. 13 was used to calculate how the conductivity of a rock containing a thin film of grain-boundary graphite varies as a function of grain size (Fig. 3). In these calculations we assumed that the graphite in the films has a random orientation and took average conductivities to be equal to those of the two natural graphites given in ref. 14. Figure 3 indicates that rocks with an average grain size of 1 cm or less will have a conductivity within the range of typical high-conductivity zones measured in the lower crust¹⁵.

Grain-boundary graphite has been postulated as a source of the high conductivity in the mantle¹⁶, but the processes by which such a film may form have not been delineated nor has the possibility been considered that such a film may occur in crustal rocks as well. The graphite film probably did not form at the temperatures and oxygen fugacities that are recorded by the silicates in the rocks, and we believe instead that it was deposited on cooling by reduction of a CO_2 -rich fluid. It is easy to understand how the grain-boundary graphite formed in the samples of Sybille Monzosyenite because these rocks crystallized at conditions of graphite-saturation and only a small amount of cooling along the buffering surface, olivine-haematite (in ilmenite)-quartz, would precipitate graphite from the fluid phase. It is not so clear, however, why graphite would have formed in the gabbroic anorthosite (LAC-10B), which probably formed originally at oxygen fugacities above those of graphite saturation. We suggest that during cooling, interoxide equilibrium caused the oxygen fugacity to approach that of graphite saturation and with further cooling graphite was precipitated, accompanied by oxyexsolution (see Fig. 4). The process of oxyexsolution, by which titanomagnetite clears itself of titanium¹⁷, can proceed only if there is a donor for oxygen. It has been argued¹¹ that one important source is the CO_2 fluid phase, by the reaction



During cooling the fluid composition would then follow the graphite saturation surface, with oxyexsolution occurring concurrently with precipitation of grain boundary graphite.

This process of graphite precipitation apparently is not restricted to the sample described here. The exsolved and reconstituted oxide compositions from the Sybille Monzosyenite and the Maloin Pluton, both of the Laramie Anorthosite Complex, plot

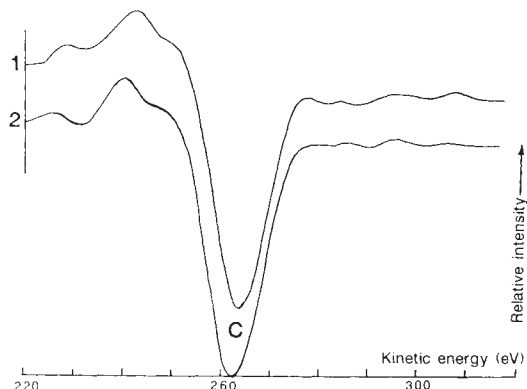


FIG. 2 A comparison of the Auger spectrum from grain-boundary graphite in LAC-8 (1) with a spectrum from Sri Lanka graphite (2). C denotes absorption due to graphite.

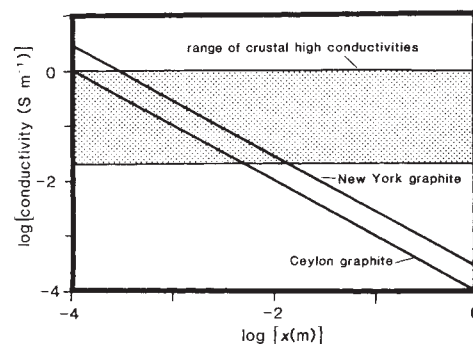


FIG. 3 Calculated conductivity for rocks that contain a 1,000-Å-thick graphite film on the grain boundaries. Calculations use the equation of ref. 13 and the average conductivities for New York and Ceylon graphite given by ref. 14. Note that rocks having an average grain size of 1 cm or less will have conductivities equivalent to those seen in high-conductivity zones in the crust.

at temperatures and oxygen fugacities compatible with graphite saturation^{11,18}. Furthermore, almost all exsolved and reconstituted oxide compositions from granulites also plot on the graphite surface¹⁹. This leads us to conclude that graphite precipitation is a common feature in high-grade rocks containing two oxides and a CO_2 -rich fluid phase, even if the rock did not have graphite as part of the primary assemblage.

We recognize three major rock types in which graphite may occur in crustal environments: (1) pelitic schists and gneisses; (2) mafic plutonic rocks; and (3) charnockitic and granulitic rocks. In graphitic pelitic rocks, it is the graphite that forms on the grain boundaries during cooling which is important for conductivity, rather than the relatively coarse individual grains seen petrographically. In most pelitic rocks, during metamorphism the fluid will be driven to the maximum water composition by continuous dehydration reactions⁷. During cooling the composition of the fluid with which graphite can coexist will be driven to higher H_2O contents, whereas the composition of the fluid in the pore spaces will be driven to compositions lower in H_2O by dehydration reactions²⁰. Both processes would tend to drive the ambient fluid composition into the graphite field, thus precipitating graphite on grain boundaries (compare Fig. 6 of ref. 8).

In rocks that crystallize at high temperatures and in the presence of CO_2 -rich fluids, such as many igneous rocks²¹ and

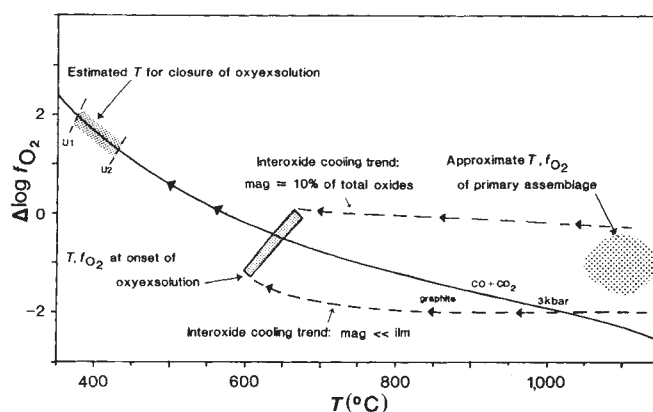


FIG. 4 Plot of T against $\Delta \log f_{\text{O}_2}$ (where $\Delta \log f_{\text{O}_2}$ is the deviation from the FMQ buffer), showing the cooling path undergone by sample LAC-10B. Original crystallization conditions were around 1,100 °C and somewhat below the FMQ buffer (as indicated by the pyroxene-QUILF equilibrium²⁸). Initially cooling followed a trend governed by Fe-Ti exchange between the oxides²⁹. Upon reaching graphite saturation, as marked by the box outlined by solid lines, oxyexsolution began, accompanied by precipitation of graphite from the fluid phase.

granulites²², a film of graphite may form on grain boundaries during cooling, even if graphite is not stable in the high-grade mineral assemblage. Two processes by which this can happen have been described above. Another process of secondary graphitization, involving precipitation of both graphite and magnesite, has been postulated to account for the formation of late graphite in the Duluth Gabbro²³.

Although often referred to as 'lower-crustal conductivity anomalies', the depth of the high-conductivity zones is variable. Indeed, high-conductivity zones are interpreted as extending nearly to the surface in some areas of Scotland²⁴. We propose that there are two causes for the increase in conductivity with depth. The first is that graphite stability is greatly enhanced by pressure. Thus a mafic rock crystallizing in a shallow environment may not reach graphite saturation, even on cooling, whereas one cooling at greater depth is more likely to develop grain-boundary graphite. The second possible cause is loss of conductivity during decompression. Grain-boundary graphite will act as a strong conductor only if the film is continuous.

During decompression one would expect that the development of fractures along the grain boundaries would destroy the continuity of the graphite film and decrease the conductivity of the rock. This has, of course, adverse implications for laboratory measurements of conductivity. One would expect that the destruction of the film is irreversible. Once the continuity of the film has been broken it will not be healed, even if the rock is subjected to high pressures. Thus it may be impossible to measure in the laboratory the conductivity that a lower-crustal rock had when it was present in the lower crust.

Our observation that grain-boundary graphite may occur in many rocks in the lower crust greatly enlarges the number of rock types that can produce high-conductivity anomalies. This provides the geophysicist with many more models from which to choose in interpreting the geophysical properties of the lower crust. More importantly, to the petrologist it obviates the necessity, when accounting for high-conductivity anomalies, of considering interlocking surfaces of brine in the lower crust^{25,26}, a situation for which evidence is largely lacking²⁷. □

Received 21 February; accepted 16 May 1989.

1. Lebedev, E. B. & Khitarov, N. I. *Geochim. Int.* **1**, 193–197 (1964).
2. Hyndman, R. D. & Hyndman, D. W. *Earth planet. Sci. Lett.* **4**, 427–432 (1968).
3. Hermance, J. F. *Geophysics* **38**, 3–13 (1973).
4. Stesky, R. M. & Brace, W. F. *J. geophys. Res.* **41**, 529–547 (1978).
5. Alabi, A. O., Camfield, P. A. & Gough, D. I. *Geophys. J. R. astr. Soc.* **43**, 815–833 (1975).
6. French, B. M. *Rev. Geophys.* **4**, 223–253 (1966).
7. Ohmoto, H. & Kerrick, D. M. *Am. J. Sci.* **277**, 1013–1044 (1977).
8. Frost, B. R. *Am. J. Sci.* **279**, 1033–1059 (1979).
9. Glassley, W. *Nature* **295**, 229–231 (1982).
10. Lamb, W. & Valley, J. *Nature* **312**, 56–58 (1984).
11. Fuhrman, M. L., Frost, B. R. & Lindsley, D. H. *J. Petrology* **29**, 699–729 (1988).
12. Frost, B. R. & Touret, J. R. L. *Contr. Miner. Petrology* (in the press).
13. Waff, H. S. *J. geophys. Res.* **79**, 4003–4010 (1974).
14. Keller, G. V. in *Handbook of Physical Constants* (ed. Clark, S. P.) 553–577 (Geological Society of America, 1966).

15. Haak, V. & Hutton, R. in *The Nature of the Lower Crust* (eds Dawson, J. B., Carswell, D. A., Hall, J. & Wedepohl, K. H.) 35–49 (Geological Society, London, 1986).
16. Duba, A. G. & Shankland, T. J. *Geophys. Res. Lett.* **9**, 1271–1274 (1982).
17. Buddington, A. F. & Lindsley, D. H. *J. Petrology* **5**, 310–357 (1964).
18. Kolker, A. & Lindsley, D. H. *Am. Miner.* **74**, 307–324 (1989).
19. Frost, B. R. & Chacko, T. *J. Geol.* (in the press).
20. Yardley, B. W. D. *Geology* **9**, 405–408 (1981).
21. Frost, B. R. & Frost, C. D. *Nature* **327**, 503–506 (1987).
22. Touret, J. *Lithos* **4**, 423–436 (1971).
23. Pasteris, J. D. *Geology* **16**, 804–807 (1988).
24. Hutton, V. R. S., Ingham, M. R. & Mbipow, E. W. *Nature* **287**, 30–33 (1980).
25. Shankland, T. J. & Ander, M. E. *J. geophys. Res.* **88**, 9475–9484 (1983).
26. Gough, D. I. *Nature* **323**, 143–147 (1986).
27. Yardley, B. W. D. *Nature* **323**, 111 (1986).
28. Frost, B. R., Lindsley, D. H., Anderson, D. J. & Davidson, P. M. *Geol. Soc. Am. Abstr. Progr.* **19**, 669 (1987).
29. Frost, B. R., Lindsley, D. H. & Andersen, D. J. *Am. Miner.* **73**, 727–740 (1988).

Imaging of deep fluids in Archaean crust

R. C. Bailey*, J. A. Craven†, J. C. Macnae‡
& B. D. Polzer‡

* Geology and Physics Departments, University of Toronto, Toronto, Canada M5S 1A7

† Physics Department, University of Toronto, Toronto, Canada M5S 1A7

‡ Lamontagne Geophysics, 24 Mowat Avenue, Toronto, Canada M6K 3E8

DESPITE the importance of aqueous fluids in tectonic, metamorphic and magmatic activity in the Earth's crust, few techniques are available for directly observing their distribution at depth, even in stable continental crust. Such observations as have been made (electromagnetic deep sounding, which by inference^{1,2} images saline fluids, and super-deep drilling³) suggest that fluids are present in varying amounts throughout the crust, in direct conflict⁴ with the normal petrological inference of a very dry deep crust. Here we report the results of an unusually high-resolution electromagnetic sounding method to image the deep electrical-conductivity structure below a 30-km profile in Ontario which crosses Archaean lower-crustal material of high metamorphic grade exposed by a Proterozoic thrust fault, the Ivanhoe Lake cataclastic zone. The electrical-resistivity image obtained shows, embedded in very resistive crystalline crust, two nearly continuous, sub-horizontal zones of enhanced electrical conductivity at depths of 2–3 and 5–6 km. Below this, the commonly observed^{2,5–7} mid-crustal rise in conductivity is seen at depths of 15–20 km. The image does not correlate well with the structure inferred from surface geology, suggesting that the transport and distribution of fluids was not strongly controlled by lithological boundaries or the thrust fault itself.

Figure 1 shows the location of the experiment undertaken

during the autumn of 1987 as part of the Lithoprobe project⁸ to investigate the structure of the Archaean crust in the Superior province near Kapuskasing, Canada. Previous electromagnetic sounding (EM) studies (9, 10 and R. D. Kurtz *et al.* in preparation) across the block have determined that the electrical properties of the region, but for a variable overburden, are roughly plane-layered to at least a depth of 25 km. The EM system chosen for this particular study was the UTEM controlled-source ungrounded-loop time-domain system¹¹. This particular system measures the step response of the ground, which has benefits for deep-penetration measurements¹². These previous EM results justified our processing of the UTEM data using the quasi-layered-Earth approach of Macnae and Lamontagne¹³. Time-domain response data in 20 channels were collected at 100-m intervals at each of about 300 receiver sites for, in most cases, the nearest three transmitter-loop positions, for a total of about 1,000 distinct receiver-transmitter pairs. The fitting process used a weighted average of all data collected within 2 km of each station. Because the EM source is inductive, only horizontal current flow is induced in any half-space or horizontally layered earth, and thus estimates are of horizontal conductivity. Because of the averaging and the diffuse nature of the surface magnetic fields, lateral location of horizontal features and vertical resolution are no better than several hundred metres (at the surface) or ~30% of the fitted depth (whichever is the greater).

Figure 2a shows the electrical conductivity section over a horizontal distance of ~30 km. The range of the data is not sufficient to determine the bottom of the region of steeply rising conductivity that occurs between 15 and 20 km. Most of the crust is extremely resistive, with resistivities of the order of 10⁵ Ω m or higher; superimposed on this are a number of features with higher conductivities (green to red). These features are irregular, and some of the irregularities of the deeper ones may be partly due to imperfect correction for irregular surface over-

burden. The shallowest features are an irregular set of near-surface conductors (yellow to red). Two processing approximations, namely that speed-of-light time-delays are zero and that the horizontal conductors are infinitely extended, have caused these features to plot at depths of up to a few hundred metres rather than at the surface. In this region, Abitibi clay-belt deposits have been mapped¹⁴, and, assuming a typical conductivity of $0.04\text{--}0.02\text{ S m}^{-1}$, the conductances imply that irregular clays and tills of thickness from zero to 2–25 m are present under the survey line. With a few exceptions, the upper 15 km of crust is almost perfectly resistive (blue). The first feature of interest at depth is a set of narrow disconnected conductors (green) appearing at a depth of ~ 2 km. There is a small one to the west, and an extensive central one apparently extending to surface. Because our processing, like that of conventional seismic reflection, assumes shallow dips, high-dip features such as this are not rendered reliably. Second, another set of narrow conductive features (green) at ~ 5 km form a quasi-continuous band across the section, with a gap approximately in the vicinity of the (here imperfectly defined) Ivanhoe Lake cataclastic zone (ILCZ). Third, there is a mid-crustal conductive feature (green) below 15 km. With the extremely diffuse current system expected at depths below 15 km, little lateral resolution can be expected across the profile at this depth and apparent deviations of the resistivity contours from the horizontal are probably the result of data errors.

The widespread existence of brines in this part of the Canadian Precambrian Shield¹⁵ supports the generally accepted attribu-

tion of crustal conductivity enhancements to saline waters¹. The most striking feature of the conductivity section is that the deviations from a very resistive background occur in fairly well defined nearly continuous subhorizontal zones. To obtain more quantitative estimates, we performed stabilized inversions¹⁶ by fitting to one-dimensional models the raw waveform data from sites in the middle of the western (left) half of the profile. A typical result is shown in Fig. 2b. The zone thicknesses obtained in the inversion are 0.5 and 1.0 km respectively, but these are probably best regarded as upper bounds because of the limited

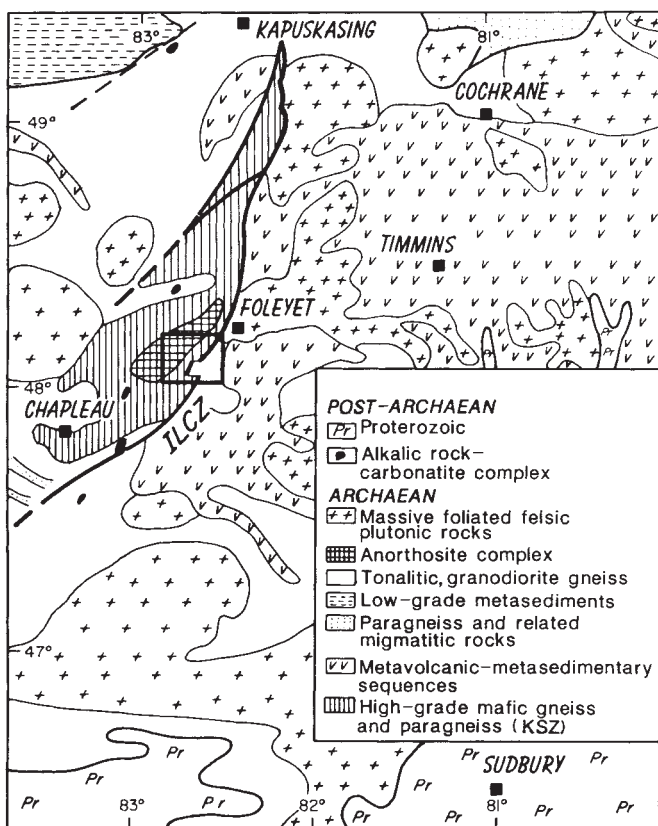


FIG. 1 Map showing the UTEM profile location (heavy rectangle near Foleyet). The Kapuskasing Uplift¹⁸ corresponds to the eastward transition of low-grade metamorphic rocks of the Wawa and Quetico belts to the high-grade amphibolites and granulites of the Kapuskasing Structural Zone (KSZ) and is interpreted as an oblique cross-section of Archaean crust exposing ~ 20 km of vertical succession¹⁷. Delimiting the eastward extent of the KSZ is the Ivanhoe Lake cataclastic zone (ILCZ)¹⁷ which is thought to represent the surface expression of an east-verging, listric thrust ramp active during the Early Proterozoic era¹⁹.

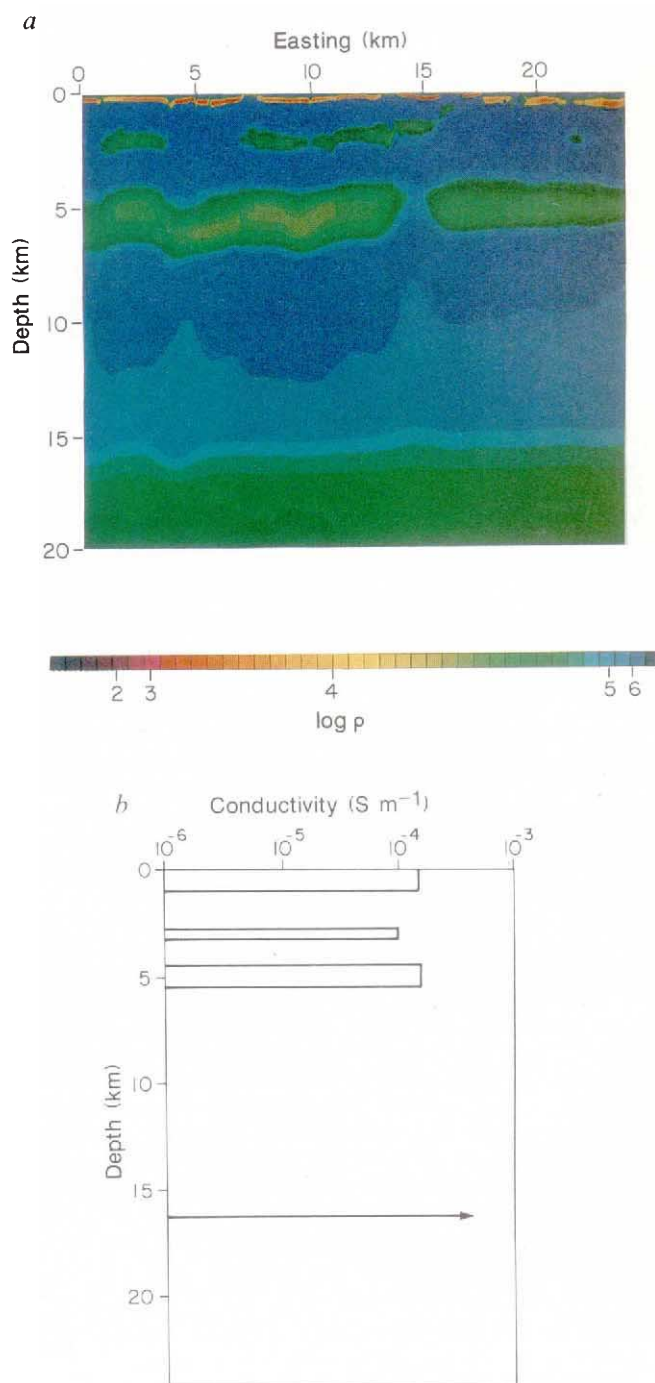


FIG. 2 a, Colour section showing log (base 10) of electrical resistivity in Ωm . West is on the left. All receiver stations have been projected onto an East-West line. The scale relating to resistivity is quite nonlinear. Highway 101 in Fig. 3 is at easting 15.3 km. b, Inverted layered conductivity model for data from the central part of the western half of the profile. Conductivity σ in S m^{-1} is plotted against depth in m. The number of layers was implicitly fixed *a priori*. Depth scale is the same as a for ease of comparison.

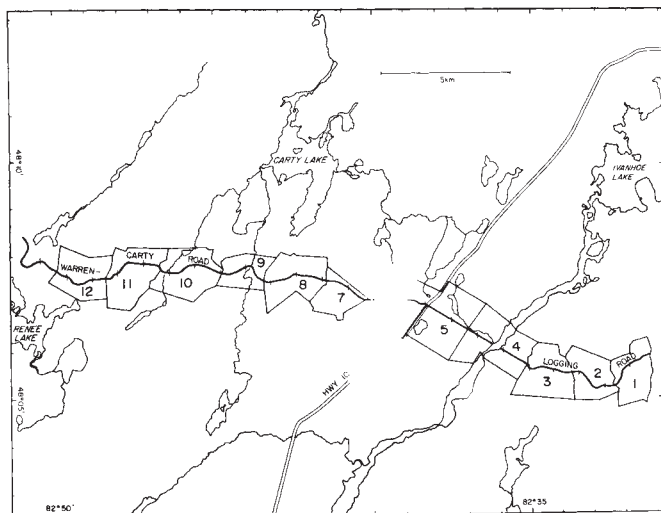


FIG. 3 Detail map showing the 12 ungrounded transmitter loop locations (solid lines) and the receiver survey line (heavy solid line). The area covered corresponds to the heavy rectangle in Fig. 1.

resolving power of the method. From these, upper bounds on the resistivity in these two conductive zones must lie in the range 5,000–10,000 Ω m, at least an order of magnitude less than the average country rock. Application of Archie's law with an exponent of 2 and water resistivity of 0.2 Ω m (representative of brines) yields a lower bound on connected porosity of $\sim 0.5\%$ for the two conductive zones at 2 and 5 km.

The surface geology is interpreted^{17–19} as the uplift of high-grade metamorphic rocks along a north-west dipping crustal-scale thrust fault, the surface expression of which is the ILCZ (approximately centred on Highway 101 in Fig. 2). Preliminary results²⁰ of the Lithoprobe deep-seismic-reflection survey along the same line show westward-dipping (~ 20 – 30°) reflectors in general agreement with the thrust-fault interpretation¹⁷ of the ILCZ. The conductive zones, on the other hand, are subhorizontal. This implies that lithological boundaries have not been the primary controls on the transport and distribution of water here, and that porosity creation in these zones post-dates the thrusting event. It seems plausible that storage of water may be made possible by horizontal hydraulic fracture zones arising from a compressive state of stress in the crust. On the other hand, any water with access to the dipping conduits that might be associated with the fault would escape, leading to the observed disruptions in the two conductive zones at the fault location. An alternative explanation of the electrical-conductivity results is that subhorizontal lithological boundaries, faults or detachment zones may exist which do not have a surface expression. This explanation might reasonably apply to the 2-km conductor, which would then represent a much more shallowly dipping geological structure (lithological unit, fault or detachment zone) than is inferred from the surface geology; it seems less plausible as an explanation for the extensive and more horizontally continuous 5-km conductor. □

Received 9 February; accepted 5 June 1989.

- Shankland, T. J. & Ander, M. E. *J. geophys. Res.* **88**, 9475 (1983).
- Gough, D. I. *Nature* **323**, 143–144 (1986).
- Lavrusenko, I. P., Byeleviski, G. A., Prizov, Y. Y. & Stolyarov, D. E. *The Superdeep Well of the Kola Peninsula* (Springer, Berlin, 1986).
- Yardley, B. W. D. *Nature* **323**, 111 (1986).
- Thompson, B. G., Nekut, A. & Kuckes, A. F. *J. geophys. Res.* **88**, 9461–9473 (1981).
- Haak, V. & Hutton, V. R. S. *Geol. Soc. Lond. Spec. Publ.* **24**, 35–49 (1986).
- Jones, A. G. *Geophys. J.* **89**, 7–18 (1987).
- Clowes, R. N. *Geosci. Can.* **11**, 122–126 (1984).
- Woods, D. V. & Allard, M. A. *Phys. Earth planet. Inter.* **42**, 135–142 (1986).
- Kurtz, R. D., Macnae, J. C. & West, G. F. *Geophys. J.* (in the press).
- West, G. F., Macnae, J. C. & Lamontagne, Y. *Geophysics* **49**, 1010 (1984).
- Spies, B. R. *Geophysics* (in the press).
- Macnae, J. C. & Lamontagne, Y. *Geophysics* **52**, 545–554 (1987).
- Boissoneau, A. N. *Ontario Geological Survey Map S365* (1965).
- Frape, S. K. & Fritz, P. in *Saline Water and Gases in Crystalline Rocks*, *Geol. Ass. Can. Spec. Pap.* **33**, 19–38 (1987).
- Potzer, B. D. thesis Univ. of Toronto (1985).
- Percival, J. A. & Card, K. D. *Geology* **11**, 323–326 (1983).
- Percival, J. A. & McGrath, P. H. *Tectonics* **5**, 553–572 (1986).
- Percival, J. A. & Krogh, T. E. *Can. J. Earth Sci.* **20**, 830–843 (1983).
- Green, A. et al. *Tectonophysics* (in the press).

Exceptionally well preserved pterosaur wing membrane from the Cretaceous of Brazil

David M. Martill* & David M. Unwin†

* The Open University, Department of Earth Sciences, Milton Keynes, MK7 6AA, UK

† Department of Zoology, Reading University, PO Box 228, Reading, Berks RG6 7AJ, UK

FOSSILIZED impressions of pterosaur wing membranes are known from a number of localities^{1,2}, but true soft-tissue preservation is extremely rare^{3,4}. We present the first description of the internal anatomy of the wing membrane, based on exceptionally well preserved soft tissues from the forearm of a Lower Cretaceous, Brazilian pterosaur. A thin epidermis overlies a dermis composed successively of a 'stratum vasculosum', 'stratum spongiosum' and a layer of striated muscle. This exceptional specimen provides important new insights into pterosaur biology. Incipient wrinkles and an apparent lack of stiffening fibres suggest the proximal region of the wing membrane was thin, extensible and tensioned by the hind limb^{5,6}. Excess metabolic heat resulting from rigorous activity could be lost by vasodilation of the vascular layer, a mechanism consistent with active flight in pterosaurs^{7–9}.

The Lower Cretaceous Santana Formation (Aptian/Albian) of the Chapada do Araripe, north-east Brazil is famous for the remarkable preservation of its fossil vertebrate fauna¹⁰. Fish are the most common¹¹, but turtles, crocodiles and dinosaurs have also been reported. Pterosaurs are the most abundant tetrapod^{12,13}, although their apparent diversity, twelve species in seven genera, is probably taxonomically inflated¹⁴.

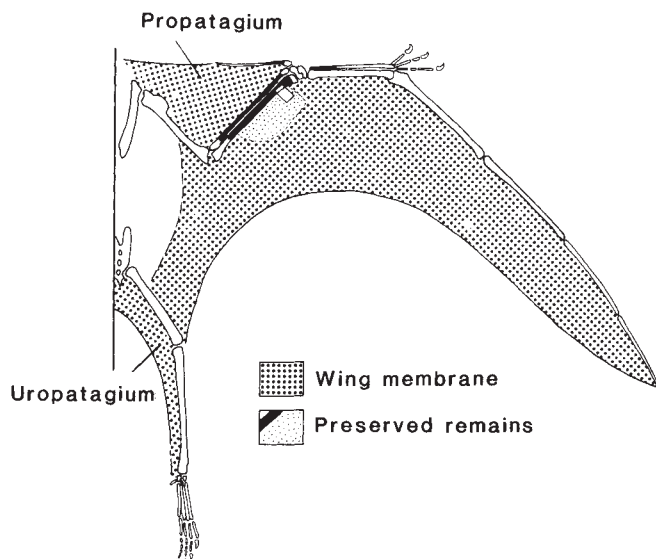


FIG. 1 Schematic reconstruction of right side of pterosaur indicating preserved remains of DNPM D6M 488 LE, and location of sample studied. Based partly on Wellnhofer^{12,13}.

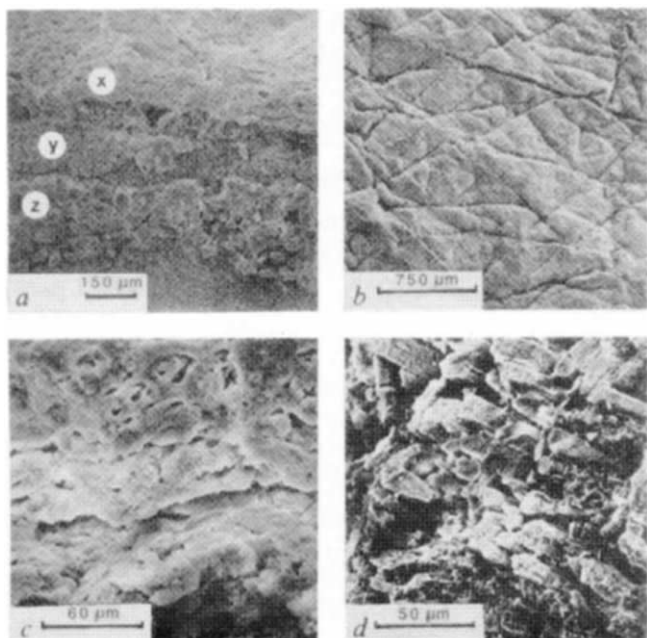


FIG. 2 Scanning electron micrographs of phosphatized pterosaur wing membrane. *a*, Acid etched vertical section through wing membrane. An upper vesicular layer 'stratum vasculosum' 'x' can be seen above an amorphous layer 'y' and a muscular layer 'z'. *b*, Detail of surface layer showing complex criss-cross pattern of epidermis. *c*, Detail of vesicular layer, here interpreted as stratum vasculosum. *d*, Detail of polygonal muscle fibres of the lowermost preserved layer.

The first example of pterosaur soft tissue preservation from this locality was reported in 1984 (ref. 4), but no details of the histology were presented. The specimen described here was removed from a small concretion (Department Nacional Produccion Mines (DNPM), Rio de Janeiro, specimen number D6M 488-LE) containing a number of incomplete but associated wing bones, representing an unnamed pterodactyloid pterosaur of approximately 2.5-m wingspan, and thus of similar dimensions to *Santandactylus pricei*¹². The sample, from a point about 10 cm distal to the elbow (Fig. 1), consists of a small tablet of rock (20×45 mm), bounded anteriorly by a section of ulna, and bearing the preserved wing membrane on its surface. A second surface, cut orthogonally to the bedding shows the ulna and

wing membrane in transverse section (Fig. 2*a*). After partial etching in 10% acetic acid for 5 min, the specimen was silver-coated¹⁵ and examined by scanning electron microscopy.

The membrane surface adjacent to the ulna appears smooth to the naked eye, but this gives way distally to a series of en echelon, low, rounded ridges (Fig. 2*b*) between 3 and 8 mm long, and oriented perpendicular to the ulna. The epidermis consists of a thin 'stratum corneum' ~0.02 mm thick (Figs 2*a* and 3*a*). Its surface is strikingly similar to the glabrous skin of humans, with criss-cross pattern consisting of irregular polygons bounded by deep furrows (Figs 2*b* and 3). Contrary to the typical reptilian condition, the epidermis is relatively thin, as in birds¹⁶ and bats¹⁷.

The epidermis merges into the dermis, the uppermost region of which consists of a poorly stratified layer of variable thickness (0.1–0.25 mm), containing numerous irregularly shaped vesicles (Fig. 2*c*). The structures probably represent the remains of sinus capillaries; they are of a similar morphology, distribution and relative size to those found in stratum vasculosum of *Gallus*¹⁸, to which this layer probably corresponds.

A layer beneath the stratum vasculosum contains some amorphous organic material, but is largely composed of coarsely crystalline calcite (Figs 2*a* and 3). It varies in thickness from 0.05 to 0.4 mm, and may correspond to the stratum spongiosum of the dermis¹⁹, typically a network of collagenous and elastic fibres set in an amorphous groundmass. This tissue is usually less robust than adjoining tissues; here it appears to have decayed more rapidly, the remaining void becoming filled with diagenetic calcite.

The stratum spongiosum is underlain by a sheet of striated muscle, varying from 0.05 to 0.45 mm in thickness (Fig. 2*a*). Individual fibres are sub-parallel to the long axis of the ulna, have polygonal cross-section, and are ~0.04 mm in maximum diameter. Transverse striae and nuclei are visible at high magnifications. The lower surface of the muscle sheet is in direct contact with the sediment, the remainder of the membrane, probably consisting of the complimentary stratum spongiosum; stratum vasculosum and epidermis is missing. The preserved section is an even 0.6 mm in depth, indicating an original minimum thickness of ~1 mm, and in general agreement with previous estimates².

Besides shedding new light on the histology of the pterosaur wing membrane, the new specimen offers important insights into aspects of thermoregulation. The high demands of active flight lead to the production of excess metabolic heat, much of

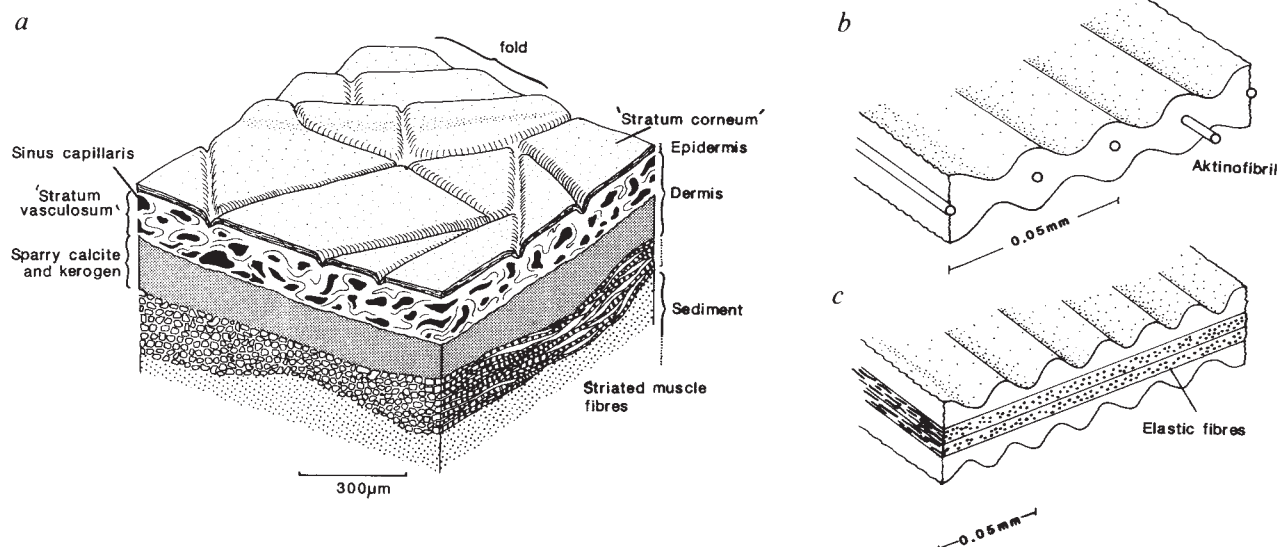


FIG. 3 *a*, Reconstruction of fossil skin sample indicating location of distinct tissue types and their relationship to sedimentologic and diagenetic fabrics. Based on the sample described here. *b*, Reconstruction of wing membrane

based on Wellnhofer². *c*, Reconstruction of wing membrane based on Pennycuik⁶.

which is lost through the wings in bats²⁰. A similar role for pterosaur wings, in which vasodilation of the stratum vasculosum permitted large-scale heat loss through the thin epidermis seems plausible. In addition, this mechanism is consistent with the idea that pterosaurs were active fliers rather than gliders⁷⁻⁹.

The structure of the pterosaur wing remains controversial. Two models have emerged, both based on wing impressions from the Solnhofen Limestone (Upper Jurassic), which show numerous, fine, regular striations^{2,8,21,22}. Wellnhofer^{8,23} and Padian⁹ have interpreted these as resulting from a wing consisting of a fibre-filled membrane (the 'aktinopatagium' of Schaller²⁴; see Fig. 3b), which was relatively stiff from front to back, but transversely extensible^{2,9}. Pennyquick^{5,6}, following Bramwell and Whitfield²⁵, interpreted the same impressions as small, regular wrinkles in a thin, highly extensible membrane (Fig. 3c), containing elastic fibres, but lacking any inherent stiffness.

The new specimen does not allow us to completely resolve this controversy. The preserved membrane contains no stiffening fibres, nor is there any evidence that they were preserved at a deeper level. Ridges seen on the surface of the epidermis of the new specimen would appear to represent wrinkles or folds, and may be of taphonomic origin. Striations have been observed adjacent to the forearm in some Solnhofen pterosaurs^{2,8,22}, but are elongated, thin parallel structures, unlike those described here. These differences may reflect variations in wing structure; moreover, the Solnhofen pterosaurs are only distantly related to the Brazilian pterosaurs, and are considerably smaller.

There is evidence for a variably stiffened wing membrane in the Upper Jurassic rhamphorhynchoid *Sordes pilosus*³. In this species the wing membrane shows a corrugated texture in the wingfinger region, but is smoother towards the body, suggesting that the wing comprised a stiffened distal section, grading into a more elastic, fibreless proximal region. Such a pattern of internal structure could account for the lack of 'stiffening fibres' in the specimen described here, and provides an explanation for wing attachment to the hind limbs in at least some pterosaurs², including *Sordes pilosus*³. In these pterosaurs the extensible proximal region of the wing was tensioned by the rear limbs, analogous in some respects to the situation in bats, but not comparable with birds, where the continuously stiffened wing would have been severely hampered by such an arrangement. □

Received 1 March; accepted 5 June 1989.

- Wellnhofer, P. In *Handbuch der Paläoherpetologie* (eds Kuhn, O. & Wellnhofer, P.), 1-82 (Fischer, Stuttgart, 1978).
- Wellnhofer, P. *Mitt. Bayer. Staatsslg. Palaont. hist. Geol.* **27**, 175-186 (1987).
- Sharov, A. G. *Akad. Nauk. SSR, Trudy Palaont. Inst.* **130**, 105-113 (1971).
- Campos, D. A., Ligabue, G. & Taquet, P. *3rd Symp. Mesozoic Terrestrial Ecosystems, Tübingen* 37-39 (1984).
- Pennyquick, C. J. *Mems. Californian Acad. Sci.* **8**, 83-98. (1986).
- Pennyquick, C. J. *Biol. Rev.* **63**, 299-331 (1988).
- Wellnhofer, P. *Abh. Bayer. Akad. Wiss.* **141**, 1-133 (1970).
- Wellnhofer, P. *Palaontographica* **A149**, 1-30 (1975).
- Padian, K. *Palaeobiology* **9**, 218-239 (1983).
- Martill, D. M. *Palaeontology* **31**, 1-18 (1988).
- Maisey, J. G. *Trop. Fish Hobby* 26-36 (May, 1986).
- Wellnhofer, P. *Palaontographica* **A187**, 105-182 (1985).
- Wellnhofer, P. *Ann. Naturhist. Mus. Wien* **88**, 149-162 (1987).
- Unwin, D. M. *Nature* **332**, 398-399 (1988).
- Mills, A. A. *Scanning Microsc.* **2**, 1265-1271 (1988).
- Spearman, R. I. C. & Hardy, J. A. *Form Function Birds* **3**, 1-56 (1985).
- Quay, W. B. In *Biology of Bats* 2 (ed. Vratsat, W. A.) 1-56 (Academic, New York, 1970).
- Bohm, R. *Acta Univ. Agr. Brno Fac. Vet.* **10**, 95-100 (1962).
- Hildebrand, M. *Analysis of Vertebrate Structure* (3rd edn) (Wiley, New York, 1988).
- Thomas, S. P. In *Recent Advances in the Study of Bats* (eds Brock-Fenton, M., Racey, P. & Rayner, J. M. V.) 75-99 (Cambridge University Press, Cambridge, 1987).
- Zittel, K. A. *Palaontographica* **29**, 47-80 (1882).
- Wellnhofer, P. *Solnhofener Plattenkalk, Urvogel und Flugsaurier* (Solnhofen Aktien-Verein, Maxberg, 1983).
- Wellnhofer, P. *Natur Mus., Frankf.* **112**, 278-291 (1982).
- Schaller, D. *Proc. int. Archaeopteryx Conf. Eichstatt, 1984* 333-348 (1985).
- Bramwell, C. D. & Whitfield, G. R. *Phil. Trans. R. Soc.* **B267**, 503-592 (1984).

ACKNOWLEDGEMENTS. We thank D. Campos, A. Kellner and P. Brito of DNPM Rio de Janeiro. Valuable contributions were made by J. M. V. Rayner, L. B. Halstead, N. N. Bakhrina, R. G. Clements and D. Norman. Use of facilities at Leicester University, Department of Geology, and Reading University, Department of Zoology, are gratefully acknowledged. This research was supported by an Open University grant awarded to Dave Martill.

Retinoic acid causes an anteroposterior transformation in the developing central nervous system

A. J. Durston, J. P. M. Timmermans, W. J. Hage, H. F. J. Hendriks*, N. J. de Vries, M. Heideveld & P. D. Nieuwkoop

Hubrecht Laboratory, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

* TNO Institute for Experimental Gerontology, Lange Kleiweg 151, 2280 HV Rijswijk, The Netherlands

ALL-trans retinoic acid (RA) is well known as a biologically active form of vitamin A and a teratogen¹⁻⁸. The identification of nuclear receptors for this ligand⁹⁻¹¹ suggests strongly that it is an endogenous signal molecule, and measurements of RA¹² and teratogenic manipulations¹³⁻¹⁵ suggest further that RA is a morphogen specifying the anteroposterior axis during limb development. Besides the limb, RA and other retinoids affect development of other organs, including the central nervous system (CNS)^{4,6,7,16-18}. None of these other effects has been investigated in detail. Our purpose here was to begin analysing the effects of RA on CNS development in *Xenopus laevis*. We find that RA acts on the developing CNS, transforming anterior neural tissue to a posterior neural specification. These and other¹⁹ findings raise the possibility that RA mediates an inductive interaction regulating anteroposterior differentiation within the CNS. Following recent reports implicating transforming growth factor- β 2-like^{20,21} and fibroblast growth factor-like factors^{21,22} in mesoderm induction, this indicates that a different type of signal molecule (working through a nuclear receptor, not a plasma membrane receptor) might mediate inductive cell interactions during early embryonic development.

We found that treating early *Xenopus* embryos with RA causes microcephaly, as has been reported for retinoids in mammals and man (Fig. 1a). We used pulse treatments (30 min) to improve the change of specifically manipulating time-dependent events during development, and these worked from 10^{-7} M, with dramatic effects at higher concentrations (Fig. 1b). Retinol was $100\times$ less effective than RA, indicating that RA or an RA metabolite is the effective form of vitamin A.

Our observation that a low external RA concentration induces microcephaly made it important to determine whether early *Xenopus* embryos actually contain endogenous RA. We therefore extracted retinoids from RA-sensitive stages (mixed batches of 650-1,200 embryos from stages 10-15; the period of neural induction) and used HPLC analysis to detect and quantitate RA, as described previously¹² (Fig. 2). The analysis showed that stage-10-15 embryos do, indeed, contain endogenous RA, at a mean concentration of 1.5×10^{-7} M. This concentration is within the range encountered in other tissues^{12,23} and quite close to the highest RA concentration reported in the chick limb bud (5×10^{-8} M, in posterior limb bud tissue)¹².

We also used HPLC to follow the fate of ^3H -labelled RA (a 10^{-7} M, 30-min pulse; the minimum treatment for inducing microcephaly) after this was taken up from external medium by stage-10 *Xenopus* embryos. We found that the labelled RA was taken up rapidly and retained in the embryo, so that the internal [^3H]RA concentration, accumulated from the external medium, was still close to the original endogenous concentration at the time when neural transformation, a likely target for RA action (see below), begins^{24,25}. The minimum concentration of externally applied RA that must accumulate in a pulsed *Xenopus* embryo to induce microcephaly thus resembles the endogenous RA concentration. This finding is consistent with the idea that RA treatment interferes with an event that is

mediated by RA *in vivo*, and we note that there is a similar match between endogenous RA concentration and the minimum applied internal RA concentration needed to obtain an axis defect in the chick limb bud^{12,15}.

External RA concentrations above 10^{-7} M cause more dramatic defects. We ruled out that our most extreme treatment (a 10^{-5} M, 30-min pulse) works through a non-specific, toxic effect. Histological sections of treated embryos and explants showed no necrotic tissue. Embryos which were cut open and exposed to propidium iodide (PI) 2 h after RA treatment showed no more dead (fluorescent, PI-permeable) cells than the controls. The stage-22 CNS volume was also normal in such RA-treated embryos (Fig. 1c), again indicating that RA-induced CNS defects do not depend on cell death.

The microcephaly caused by RA is associated with a dramatic reduction of the anterior CNS, as described below. The microcephalic embryos lack most or all of the forebrain and midbrain, as well as the sense organs (eyes and nasal pits) that are induced by or formed from the forebrain.

A trivial explanation of the effect of RA on CNS development might be that it interferes with gastrulation, thus inhibiting mesoderm migration and consequently neural induction. This possibility was tested in two ways (Fig. 1b,c). First, we examined the timing of the RA effect. Figure 1b shows that RA pulses induce microcephaly if applied up until the early neurula stage (stage 14). RA is thus still effective after gastrulation is finished, indicating that it can work without inhibiting gastrulation. Second, we looked at the effect of RA treatment on the volume of the CNS. Figure 1c shows a comparison of CNS volumes (made at stage 22, just after the end of neurulation) between normal embryos and embryos in which extreme microcephaly was induced by a pulse treatment with 10^{-5} M RA at stage 10.

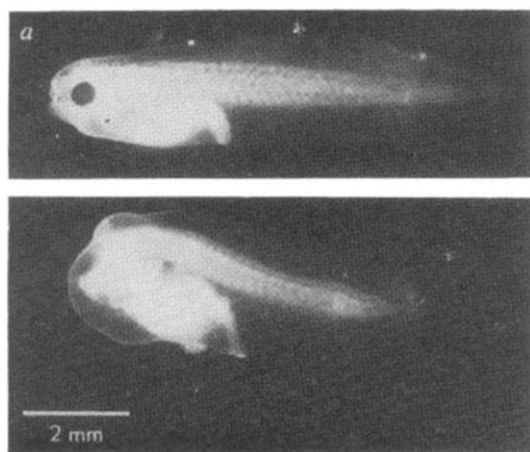
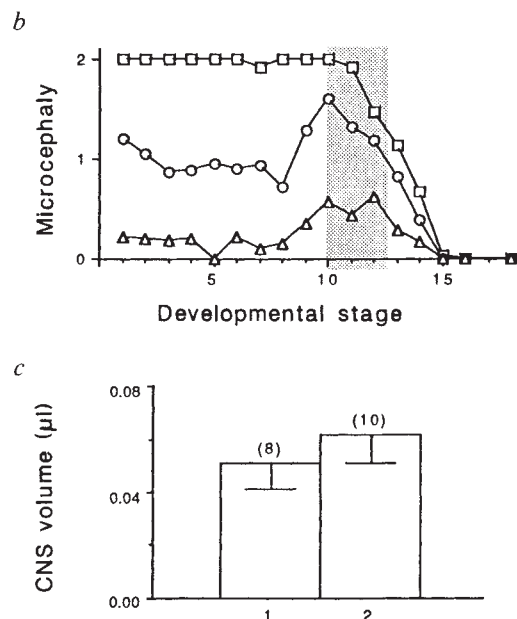


FIG. 1 *a*, RA causes microcephaly in *Xenopus*. Top: a normal *Xenopus* tadpole at stage 45. Embryos were cultured at $21 \pm 1^\circ\text{C}$ in 25% MMR (modified amphibian Ringer)³⁰. Bottom: a tadpole which developed from an embryo which had been treated with a 30-min pulse of 10^{-5} M RA (in 25% MMR), at stage 10. The embryo shows extreme microcephaly and eye formation is suppressed. *b*, The treatment-time dependence of RA-induced microcephaly. Batches of 80–100 embryos were treated with 30-min pulses of 10^{-5} M ($\square$), 10^{-6} M ($\circ$) or 10^{-7} M ($\triangle$) RA, as above, at each of the developmental stages indicated and cultured through to a late tadpole stage (stage 45), as in *a*. The tadpoles were then scored for microcephaly (2: extreme microcephaly, no eye pigment; 1: mild microcephaly, reduced eye formation; 0: normal), and a mean value was calculated for each treatment time. The figure shows that treatment with a 10^{-5} M, 10^{-6} M or 10^{-7} M RA pulse causes microcephaly if the pulse is applied at, or before, the mid-neurula stage (stage 14). Later treatments are not effective. The lower concentrations (10^{-6} and 10^{-7} M) are most effective if applied between the late blastula and neurula stages. Higher concentrations (10^{-5} and 10^{-6} M) are also effective if applied early in development, probably because RA is lipophilic and not easily washed out of embryos after a pulse treatment. The figure thus shows that microcephaly can still be induced by a pulse of

Stage 22 was chosen for analysis because at this stage the boundaries of the CNS can be determined accurately, whereas there has been very little growth of the CNS since neural plate formation. The figure shows that the CNS volume remains normal in these extremely abnormal embryos. This result also indicates that RA does not induce a drastic gastrulation defect, leading to reduced formation of the CNS. We cannot yet rule out that RA has some subtle effect on the migration or differentiation of mesoderm (see below).

An alternative idea is that RA might mimic (or might actually be) a morphogen that regulates regional differentiation of the CNS. This idea predicts that RA should exert its effects directly on the responsive neuroectoderm and that it should suppress differentiation of anterior neural structures and/or stimulate differentiation of posterior neural structures. We examined the effects of RA on neural differentiation in cultured ectoderm explants (Fig. 3a–c). Figure 3c shows that RA, at a concentration that causes extreme microcephaly, neither induces neural differentiation in competent stage-10 ectoderm nor inhibits it in competent stage-10 ectoderm that is combined with a small piece of dorsal mesoderm (that is, the natural neural inducer). This, and the fact that RA treatment of whole embryos has no effect on CNS volume at stage 22 (Fig. 1c), indicates that RA has no effect on the induction of ectoderm to neural tissue. RA does, however, affect regional differentiation within the neuroectoderm.

The neural tissue that forms in neural-induced ectoderm explants develops, after four days of culture, into structures that are recognizable as parts of the CNS. Organs formed from or induced by different CNS regions (nasal pits, eyes, ears) are also recognizable. Such regional differentiation is evident in recombinates, made from stage-10 ectoderm and dorsal mesoderm. It is also evident in neuroectoderm which has been



RA that is applied after the end of gastrulation (the gastrulation period is indicated by the hatched bar). *c*, The volume of CNS is unaffected by an extreme RA treatment. Mean volumes and standard errors for (1) 8 control and (2) 10 RA-treated embryos are shown. The CNS volume does not differ significantly between the normal and the extremely microcephalic, RA-treated embryos. Stage-10 embryos were treated for 30 min with 10^{-5} M RA (a treatment that gives extreme microcephaly in 100% of cases), and were then cultured through to stage 22 (an early stage where the extent of the CNS can be analysed unambiguously using histology). They were then fixed in Smith's fixative (2.5% acetic acid, 0.5% $\text{K}_2\text{Cr}_2\text{O}_7$, 4% formaldehyde), dehydrated, embedded in histowax, cut transversely into $10\ \mu\text{m}$ sections and stained with haematoxylin–eosin. The cross-sectional area of the CNS tissue (excluding cavities) was measured in every tenth section, using a digitizer and the CNS volume was calculated.

excised from the embryo at stage 11, after natural neural induction has occurred (although stage-11 neur ectoderm explants show mainly anterior neural differentiation). These two situations were used to explore the effects of RA treatment on regional neural differentiation. We examined the effects of a 10^{-5} M RA pulse on regional differentiation in ectoderm-mesoderm recombinates, where either the ectoderm or the mesoderm or both were treated with RA. We also examined the effects of a 10^{-5} M RA pulse on neur ectoderm that was excised at stage 11 and then treated with RA. These experiments showed a clear effect of RA on regional differentiation in the CNS; RA inhibited formation of anterior neural structures as well as sense organs (eyes, nasal pits) induced by or formed from the forebrain, but it did not block formation of posterior neural structures and ears. The suppression of anterior neural structures by RA is illustrated in Fig. 3d-f for an easily scored marker, the eye.

Figure 3f shows that eye formation is suppressed if ectoderm-mesoderm recombinates are treated with RA; also, the RA effect is dramatic if the ectoderm in the recombinant is treated with RA, but less so if the mesoderm is treated with RA. RA suppresses eye formation in excised neur ectoderm (stage 11), which was treated with RA after neural induction had occurred. These data illustrate that RA suppresses anterior neural differentiation and show that it can do this through a direct effect on the neur ectoderm, as would be expected should RA either mimic or be a morphogen, regulating regional differentiation in the CNS. We cannot yet rule out that RA also acts indirectly on CNS differentiation by altering the anteroposterior specification of mesoderm. The invaginating dorsal mesoderm normally segregates into anterior prechordal plate and posterior

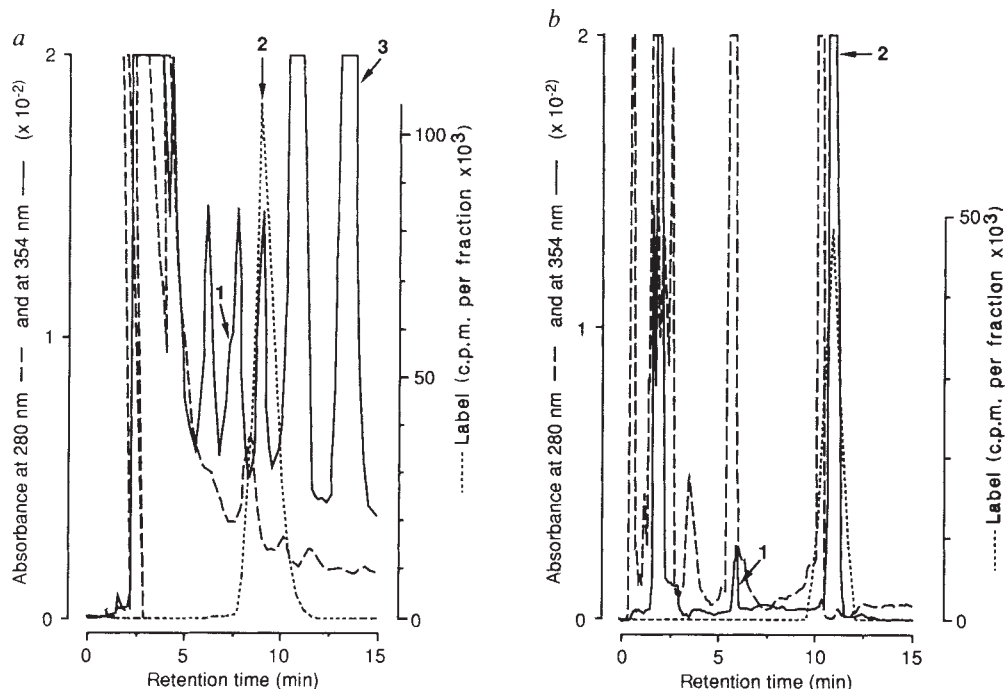
chordamesoderm. These two regions underly and induce the forebrain, and the hindbrain and spinal cord, respectively^{25,26}. Altered mesodermal specification would thus change the regional specification of the CNS. The effect of RA on anterior to posterior transitions in the CNS is dealt with in detail below.

The detailed effects of RA on CNS development can best be analysed in whole embryos because the embryo can be oriented reliably for histological analysis and landmarks defining boundaries between CNS regions can then be identified accurately, if the analysis is performed at an appropriate stage. Embryos were treated with a 10^{-5} M RA pulse at stage 10 or were untreated (controls), cultured through to stage 45, processed histologically and analysed using standard neuroanatomical markers, as described previously²⁷ (see legend to Fig. 4). The embryos treated with RA were very microcephalic, as in Fig. 1a. They lacked eyes and nasal pits but not ears. They also showed a dramatic anterior to posterior transition in the CNS. The anterior brain (forebrain and midbrain) was much reduced, but the hindbrain was increased substantially and the spinal cord was also increased (Fig. 4).

The findings above show that pulse treatments with RA, a suspected morphogen for limb development, suppress formation of the anterior CNS, if applied at an appropriate stage during early *Xenopus* development. *Xenopus* embryos contain endogenous RA, and the minimum internal concentration of externally applied RA that must accumulate in an embryo to induce microcephaly resembles the endogenous RA concentration, suggesting that RA treatment acts by interfering with an event that is mediated by RA *in vivo*. The time-dependence of the RA effect and the fact that RA has no effect on CNS volume

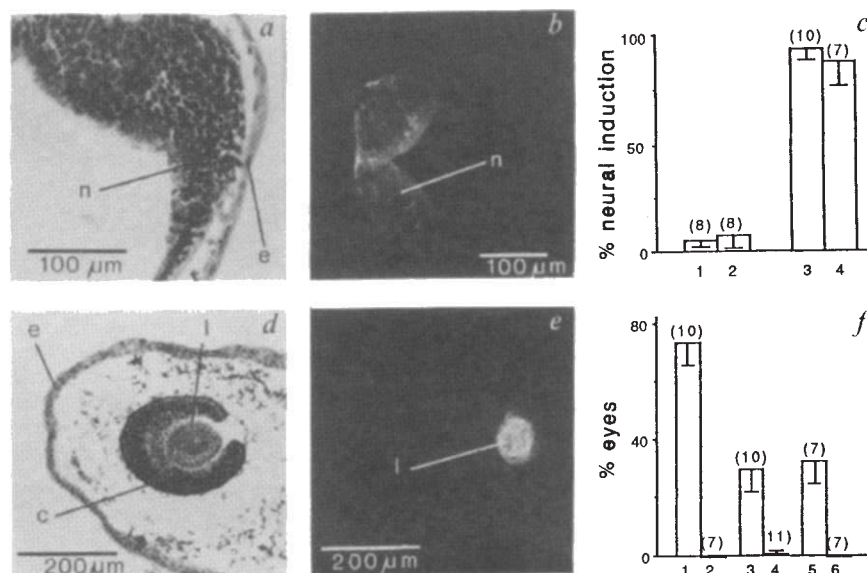
FIG. 2 The HPLC sample chromatograms presented illustrate the purification and identification of all-*trans*-retinoic acid (RA) from stage-10-15 *Xenopus laevis*. Broken lines, absorbance at 280 nm; solid lines, absorbance at 354 nm; dotted lines, elution profiles of the internal standard ^3H -RA, determined by scintillation counting. *a*, Fractionation of extracts of *Xenopus* embryos on a C18 column (solvent A at a flow rate of 1.0 ml min^{-1}) results in five major peaks that absorb strongly at 354 nm but weakly at 280 nm (typical for retinoids). The peak labelled 2 coelutes with all-*trans*- ^3H -RA. The numbers indicate the positions of external standards: 1, 13-*cis*-retinoic acid; 2, all-*trans*-retinoic acid; and 3, retinal. *b*, Fractionation of extracts of *Xenopus* embryos on a silica column (solvent B at a flow rate of 1.0 ml min^{-1}) results in two major peaks that absorb strongly at 354 nm but weakly at 280 nm (typical for retinoids). The peak labelled 2 coelutes with ^3H -RA. The numbers indicate the positions of external standards: 1, 13-*cis*-retinoic acid; 2, all-*trans*-retinoic acid.

METHODS. 650-1,200 embryos ($\sim 650\text{--}1,200\text{ }\mu\text{l}$) were collected into a vial and frozen immediately in liquid nitrogen. On thawing, $50\text{ }\mu\text{l}$ stabilizing buffer³², $100\text{ }\mu\text{l}$ saturated sodium sulphate solution, $40\text{ }\mu\text{l}$ ethyl acetate/methyl acetate (8:1) (v/v) and $40\text{ }\mu\text{l}$ ethanol were added for each $100\text{ }\mu\text{l}$ of embryo homogenate. As an internal standard, $1\text{ }\mu\text{Ci}$ of all-*trans*-(11,12- $^3\text{H}_2$)-retinoic acid (specific activity 55.6 Ci mmol^{-1}) was added for each $100\text{ }\mu\text{l}$. The tissue was sonicated on ice and extracted three times with 2 ml ethyl acetate/methyl acetate (8:1). After centrifugation, the organic phases were pooled and the solvent was evaporated in the dark under nitrogen at room temperature to a volume of $\sim 250\text{ }\mu\text{l}$. The solvent was then filtered through a Millex-HV4 0.45- μm filter. For HPLC, we used a Radial



Pack C18 analytical column, particle size $10\text{ }\mu\text{m}$, preset with an RCSS 18 Guard-PAK precolumn insert (a) and a $100\text{-mm} \times 3\text{-mm}$ silica column from Chrompack (b) with the following solvent systems: for *a*, methanol/acetonitrile/water/acetic acid/triethylamine (7.3:0.9:1.2:0.1:0.5), and for *b*, *n*-hexane/acetonitrile/acetic acid (0.995:0.004:0.001). The endogenous RA was quantitated by comparing its peak area with that of external (RA) and internal (retinyl acetate) standards. The endogenous RA from 3,500 embryos was purified using the C18 column, evaporated in the dark under nitrogen, brought into solution in Flickinger medium at 10^{-6} M, and used to treat 25 stage-9 *Xenopus* embryos. The treatment resulted in microcephaly (mean value 1.2 on the scale in Fig. 1b).

FIG. 3 *a-c*, RA neither stimulates nor inhibits neural differentiation. Ectoderm explants were excised surgically from stage-10 embryos. Some of them were recombined with a small piece (~10% of the volume of the explant) of dorsal marginal zone mesoderm, the natural neural inducer, and the explants and recombinates were either treated with a 30-min pulse of 10^{-5} M RA at the beginning of the experiment or were untreated (controls). They were then cultured for 2-4 days in 100% Flickinger solution³² before being tested for neural differentiation. *a*, The explants were usually processed histologically, using the same procedure as for whole embryos (Fig. 1, legend). Neural tissue was then recognized by its characteristic appearance after haematoxylin-eosin staining (small dense cells). Key: n, neural tissue; e, epidermis. *b*, The identity of the neural tissue was checked, in some cases by FITC immunofluorescent staining using a neural specific antibody, 2G9 (ref. 33). Key: n = neural tissue. *c*, Ectoderm explants (1,2) showed almost no neural differentiation (~5% of cases, because of excision error). This was true, whether they were treated with retinoic acid (2), or not (1). Ectoderm/mesoderm recombinates (3,4) on the other hand, showed neural differentiation in almost 100% of cases. This was true whether they were treated with retinoic acid (4) or not (3). The data are based on *n* batches of 20 explants, where *n* is the number above each histogram bar. *d-f*, RA suppresses formation of an anterior neural structure, the eye. Explants and recombinates showing neural differentiation can develop recognizable neural structures after 4 days of culture. The eye, an easily recognizable anterior neural structure, was used to assay the effect of RA pulses (as above) on anterior neural differentiation. *d*, Eyes were usually recognized by their characteristic morphology after haematoxylin-eosin histology. Key: e, epidermis; c, eyecup; l, lens. *e*, In some cases their identity was checked, using immunostaining with an antibody against total *Xenopus* lens proteins³⁴.



Key l, lens. *f*, Ectoderm-mesoderm recombinates (1), produced as in Fig. 2c, formed eyes in ~75% of cases. Treatment of these recombinates with RA (2) abolished eye formation, although the recombinates still made as much neural tissue as controls (see Fig. 2c; 3,4). If the mesoderm only in these recombinates was treated with RA, eye formation was reduced (3). If only the ectoderm was treated with RA, eye formation was abolished (4). Neurectoderm explants were also excised from stage-11 embryos, after natural neural induction had occurred. These explants formed eyes in ~35% of cases (5). Treatment with RA abolished eye formation by these neurectoderm explants (6). The data are based on *n* batches of 20 explants; *n* is given above each histogram bar.

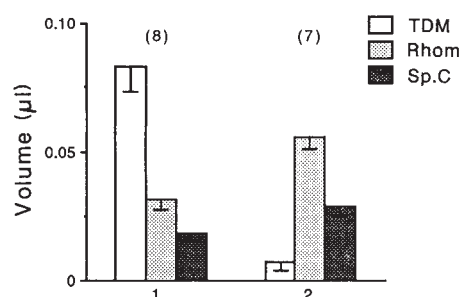


FIG. 4 The effect of RA on the composition of the CNS. Embryos were either (1) untreated (controls) or (2) were treated with a 30-min pulse of RA at stage 10. They were then cultured through to stage 45, as in Fig. 1. This is a stage at which the regions of the CNS can be distinguished from each other and at which the boundaries between the main CNS regions can be determined accurately, using morphological landmarks²⁷. The embryos were fixed, sectioned transversely and analysed by digitizing sections (see legend to Fig. 1c). The boundary between the mesencephalon and the rhombencephalon was identified by the position of the commissura cerebellaris and the boundary between the rhombencephalon and the spinal cord was identified by the disappearance of the thin roof of the rhombencephalon. These boundaries were then used to enable us to calculate the volumes of the anterior brain (telencephalon, diencephalon, mesencephalon: TDM); of the hindbrain (rhombencephalon: Rhom) and of the spinal cord (Sp. C). The figure shows mean values and standard errors for the volumes of these CNS regions from eight control embryos and seven embryos treated with RA. It is evident that the anterior-brain volume is reduced dramatically in embryos treated with RA, whereas spinal-cord volume is increased slightly and the hindbrain volume is increased substantially. It is also evident that the total volume of the CNS grows between stage 22 (Fig. 1c) and stage 45 (this figure). This growth is less in embryos treated with RA than in controls, leading to a reduction in total CNS volume in embryos treated with RA at stage 45 of ~30%. This total volume reduction is only about one half of the (absolute) volume reduction in the forebrain in embryos treated with RA. It cannot account for this, nor for the RA-induced increases in volume in the rhombencephalon and spinal cord.

show that this does not depend on a drastic inhibition of gastrulation. Nor is it due to an inhibitory (or stimulatory) effect of RA on differentiation of ectoderm to neural tissue. RA acts directly on neurectoderm to inhibit formation of anterior neural structures, and analysis of the effect on whole embryos indicates that it switches anterior-specified neurectoderm to a posterior specification. We also examined the RA effect at the molecular level, and found that a homoeobox gene messenger RNA which is only expressed posteriorly in the *Xenopus* larva shows substantially increased expression at stage 30 following a pulse treatment of the embryos with RA at stage 8 (M.H. and A.D., unpublished data).

These findings indicate that RA could be acting as an endogenous morphogen regulating regional differentiation of the CNS. In fact, they are relevant to a hypothesis²⁷⁻²⁹ that neural induction occurs in two steps. A first step, neural activation, converts competent ectoderm to anterior-specified neural tissue. A second step, neural transformation, converts anterior-specified neural tissue to a posterior specification. RA thus behaves as if it affects neural transformation but not neural activation.

The timing of the RA effect (Fig. 1b) and the direct effect of RA on neurectoderm suggest that RA influences neural transformation directly. It may also affect other processes, notably anteroposterior specification of the mesoderm. Injected ³H-labelled RA accumulates in the CNS during early mouse development and the minimum (intra-embryonic) concentration of externally applied RA that *Xenopus* embryos need to take up to induce anteroposterior transitions in the CNS is similar to their endogenous RA concentration during neural induction. This minimum externally applied concentration is also quite similar to the endogenous RA concentrations that probably specify anteroposterior axis formation in the chick limb¹². We think it worthwhile, therefore, to investigate whether RA has a direct or indirect role in specifying the anteroposterior axis of the CNS. □

Received 8 May; accepted 24 May 1989.

1. Wolf, G. *Physiol. Rev.* **64**, 3, 873-932 (1984).
2. Madani, K. A. *Nutrition Res.* **6**, 107-123 (1986).
3. Chytil, F. *J. Am. Acad. Dermatol.* **15**, 741-747 (1986).
4. Yasuda, Y. *Teratology* **34**, 37-49 (1986).
5. Sporn, M. B. & Roberts, A. B. (eds) *Retinoids, Differentiation and Disease* (CIBA Symposium 113, Pitman, London, 1985).
6. Shenfelt, R. E. *Teratology* **5**, 1, 103-118 (1972).
7. Campbell, L. R., Hayton, D. H. & Sohal, G. S. *Teratology* **34**, 171-197 (1986).
8. Fantel, A. G., Shephard, T. H., Newell-Morris, L. L. & Moffet, B. C. *Teratology* **15**, 65-72 (1977).
9. Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. *Nature* **330**, 444-450 (1987).
10. Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. *Nature* **330**, 624-629 (1987).
11. Brand, N. *et al. Nature* **332**, 850-853 (1988).
12. Thaller, C. & Eichele, G. *Nature* **327**, 625-628 (1987).
13. Tickle, C., Alberts, B. M., Wolpert, L. & Lee, J. *Nature* **296**, 564-565 (1982).
14. Summerbell, D. *J. Embryol. exp. Morph.* **78**, 269-289 (1983).
15. Tickle, C., Lee, J. & Eichele, G. *Dev. Biol.* **109**, 82-95 (1985).
16. Lammer, E. J. *et al. New Engl. J. Med.* **313**, 837-841 (1985).
17. Kocher, D. M. *Acta path. microbiol. scand.* **70**, 398-404 (1967).
18. Morriss, G. M. *J. Anat.* **113**, 241-250 (1972).
19. Dencker, L., d'Argy, R., Danielsson, B. R. G., Ghantous, H. & Sperter, G. O. *Dev. Pharmac. Ther.* **10**, 212-223 (1987).
20. Rosa, F. *et al. Science* **239**, 783-785 (1988).
21. Kimmelman, D. & Kirschner, M. *Cell* **51**, 869-877 (1987).
22. Slack, J. M. W., Darlington, B. G., Heath, J. K. & Godsave, S. F. *Nature* **326**, 197-200 (1987).
23. McCormick, A. M. & Napoli, J. L. *J. biol. Chem.* **257**, 1730-1735 (1982).
24. Eyal-Giladi, H. *Extrait Arch. Biol. LXX*, 2, 180-259 (1954).
25. Spemann, H. W. *Wilhelm Roux Arch. Entw. Mech. Org.* **123**, 389-517 (1931).
26. Sala, M. *Proc. K. ned. Acad. Wet. Sciences C58*, 635-647 (1955).
27. Nieuwkoop, P. D. *et al. J. exp. Zool.* **120**, 1-108 (1952).
28. Nieuwkoop, P. D. *Adv. Morphogenesis* **10**, 2-39 (1973).
29. Toivonen, S. & Saxen, L. *Science* **159**, 539-541 (1986).
30. Naspoa, J. W. & Kirschner, M. W. *Cell* **37**, 731-742 (1984).
31. Eichele, G., Tickle, C. & Alberts, B. *J. Cell Biol.* **101**, 1913-1920 (1985).
32. Flickinger, R. A. *J. exp. Zool.* **112**, 165-185 (1949).
33. Woodland, H. R. & Jones, E. A. *Development* **101**, 925-930 (1987).
34. McDevitt, D. S. & Brahma, S. K. *J. exp. Zool.* **186**, 127-140 (1975).

ACKNOWLEDGEMENTS. We thank Dr E. A. Jones for the use of monoclonal antibody 2G9 and Dr S. K. Brahma for the use of the anti-*Xenopus* lens antibody. We also thank S. W. de Laat, S. K. Brahma and A. P. Otte for comments on the manuscript, and F. Vervoordeldonk and W. Leurink for illustrations.

A new anatomy of the prestalk zone in *Dictyostelium*

K. A. Jermyn, K. T. I. Duffy & J. G. Williams

Imperial Cancer Research Fund, Clare Hall Laboratory, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3LD, UK

THE characteristic structure of the mature *Dictyostelium* culminant is created by the regionalized cellular differentiation and directed movement of prestalk cells. The front prestalk zone of the migratory slug has previously been considered to be a homogeneous tissue. Here we demonstrate, however, the existence of multiple classes of prestalk cells located in different parts or the slug anterior. The *pDd56* and *pDd63* genes encoding closely related extracellular matrix proteins¹⁻⁴ are dependent for their expression upon DIF-1, the specific stalk-cell inducer⁵⁻⁸. We have fused the promoters of the two genes to a modified chloramphenicol acetyltransferase (*cat*) gene to produce immunologically detectable proteins which localize to the cell nucleus. These two markers define three distinct kinds of 'prestark' cells. One class, which we term 'prestark A' cells, expressed the *pDd63* gene. 'Prestark B' cells express *pDd56* and may also express the *pDd63* gene. A third class, which we term 'prestark 0' cells, expresses neither marker.

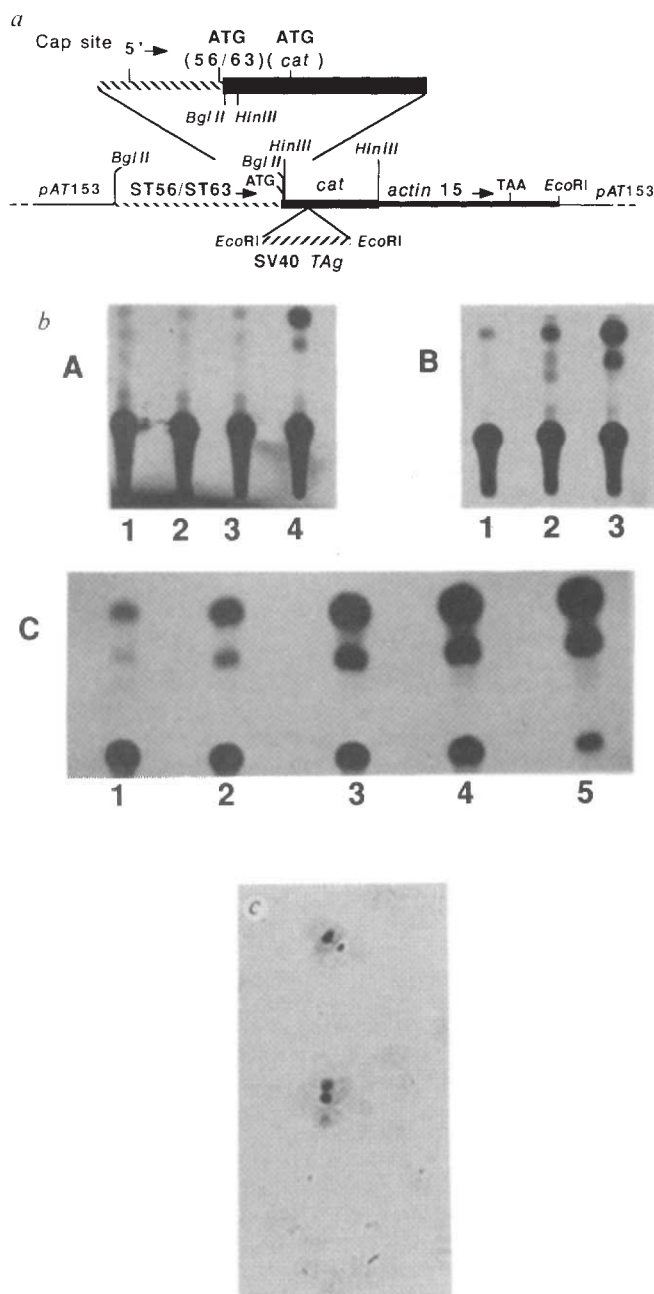
The *pDd56-Tag-cat* and *pDd63-Tag-cat* constructs direct temporally correct, cell-type-specific and DIF-inducible gene expression after transformation into *Dictyostelium* (Fig. 1a, b)^{3,9}. They contain the simian virus 40 (SV40) nuclear localization signal¹⁰, and staining of cells with monoclonal antibodies against the SV40 segment^{11,12} shows the fusion proteins to be located within the nucleus (Fig. 1c). They provide cell-autonomous markers which mirror the behaviour of the endogenous genes and can be used to establish the distribution of cells expressing them.

We have identified cells expressing the *pDd63* fusion gene in whole mounts of slugs, and find staining to be almost totally

FIG. 1 a, The structure of the *pDd56* and *pDd63* fusion gene constructs. The *pDd63* and *pDd56* genes were subjected to site-directed mutagenesis²¹ to generate unique *Bgl*II restriction sites within the sequences encoding the N-termini of their cognate proteins. The genes were then ligated as *Bgl*II fragments into the vector *pB10-actin15-cat*²² so that their upstream regions direct expression of a cytosolic Cat fusion protein. The *pDd56* fusion gene contains 2.1 kilobases of sequence upstream of the cap site; the *pDd63* gene contains 2.6 kilobases of upstream sequence. The *Bgl*II sites are positioned so that the first three amino acids of the *pDd56* protein⁹ and the first four amino acids of the *pDd63* protein² are retained within the fusion proteins. The constructs link through the *Bgl*II site, and a short polylinker in *pB10-actin15-cat*, to an in-frame *Hin*III site just upstream of the initiation codon of the *cat* gene²². This region is magnified in the upper part of the figure. These two constructs, *pDd56-cat* and *pDd63-cat*, were further modified by the insertion, at the *Eco*R1 site interrupting amino-acid 72 in the *cat* gene, of a fragment from the SV40 T-antigen (*Tag*) gene to generate *pDd56-Tag-cat* and *pDd63-Tag-cat*. To facilitate this step, the 63 and 56 *cat* cassettes were cloned into the *Bam*-*Sal* sites of the vector pAT153²³. The SV40 fragment (derived from *cat* · T · *HBsp*, see ref. 11) encodes the first 272 amino acids of the T-antigen which contains non-overlapping epitopes recognized by four different monoclonal antibodies^{10,11} and the T-antigen nuclear localization signal¹². The fusion genes direct transcription of mRNAs containing the 3' non-coding region and downstream proximal sequences of the *actin15* gene, which provide efficient termination and polyadenylation signals²⁴. b, Developmental time course and DIF-inducibility of the *pDd63-cat* fusion gene. The *pDd63-cat* construct was introduced into *Dictyostelium* cells. A single transformant clone was isolated and used to assess the regulation of *cat* activity directed by the promoter. A, Cells were collected from growth (lane 1) or subjected to development on agar and collected at the tipped aggregate stage (lane 2), the slug stage (lane 3) or the mexican-hat stage of culmination (lane 4). B, Cells were subjected to development on agar, under conditions which promote slug migration, and slugs were collected after 1 (lane 1), 2 (lane 2) and 3 days (lane 3) from initiation of development. C, Cells were subjected to development in suspension for 5 h in the presence of cAMP and then induced with DIF. Aliquots were collected at the time of DIF addition (lane 1), or after 30 (lane 2), 60 (lane 3), 90 (lane 4) or 120 min (lane 5) of induction. Constructs containing the upstream region present in *pDd56-cat* have been shown to direct temporally regulated gene expression³. The *pDd56-cat* construct itself was used in the present analyses and was found to display correct regulation (K. A. J., unpublished data). c, The *pDd56-Tag-cat* construct directs expression of a fusion protein which localizes to the nucleus. A cloned *Dictyostelium* transformant, isolated as described in the legend to Fig. 1, with a copy number for the fusion gene of ~100, was developed on agar. At the mexican-hat stage of culmination, aggregates were mechanically disrupted and individual amoebae were fixed and stained to visualize the fusion protein. The *pDd63-Tag-cat* gene directs expression of a protein which, apart from a few amino acids at the N-terminus (see legend to Fig. 1a), is almost identical in structure to that encoded by *pDd56-Tag-cat* and also localizes to the nucleus (data not shown). The characteristic, punctate, nuclear staining is useful in that it can be clearly discriminated over background, non-specific staining. Also, the level of staining is much higher than is obtained with a Cat antibody, using constructs not containing the SV40 DNA, probably because the protein is concentrated and stabilized by nuclear compartmentalization.

METHODS. The *pDd63-cat* construct was co-transformed into Ax2 cells with the vector pB10TP1²⁵ using a minor modification²⁵ of the method of Nellen *et al.*²⁶. Clones were collected from the Petri dishes in which they were initially selected, when they were ~1 mm in size, and maintained in medium containing G418 at a concentration of 10 µg ml⁻¹. A clone of transformants, in which the *pDd63-cat* gene was present at a copy number of ~100, was used in all experiments. Amoebae were developed on agar under conditions favouring or repressing slug formation²⁷. In the DIF-induction experiment, cells were starved at 10⁷ cells per ml in shaken suspension at 120 r.p.m. for 1 h. Cyclic AMP was added to a concentration of 1 mM and after a further 4 h incubation, chemically synthesized DIF-1 (a gift from Dr R. Kay) was added to a concentration of 0.1 µM. Cells were collected, washed

restricted to the front 10% of the length of the slug (Fig. 2a, b). This is a surprising result because studies using vital dyes, which stain prestalk cells more strongly than prespore cells¹³, show the first ~20% of the length of the slug to consist of prestalk cells¹⁴. We have confirmed the existence of a region which does not contain cells expressing the *pDd63* gene, between the anterior 10% and the prespore zone, by staining individual slugs



twice in 20 mM potassium phosphate buffer, pH 6.2, (KK2) and assayed for total cytosolic Cat activity²⁸. For c, transformant amoebae were developed on Millipore filters²⁹ to the mexican-hat stage and the cells were dissociated by vortexing in KK2. The cells were placed on slides coated with poly-L-lysine and allowed to settle for 10 min. They were then fixed in acetone for 2 min and incubated at 4 °C in phosphate buffered saline for 16 h with a mixture of the monoclonal antibodies PAb419 (ref. 10), 416 (ref. 10), 220 (ref. 11), and 281 (ref. 11), which bind to the SV40 sequences present in the fusion gene. The resulting immune complexes were detected by incubation for 4 h with horseradish peroxidase coupled to a goat anti-mouse antibody (Biorad). Visualization of the peroxidase was achieved by a 5 min incubation with 0.06% diaminobenzidine and hydrogen period.

in vivo with neutral red and then staining for the fusion protein ((Fig. 1c legend). About 10% of cells in the posterior region of the slug are morphologically and biochemically very similar to the prestalk cells present in the anterior region, and hence are termed 'anterior-like' cells.¹⁴⁻¹⁶ There are scattered nuclei throughout the prespore zone which are stained and we assume these to be anterior-like cells. The 'rearguard' cells, a small

FIG. 2 Analysis of the distribution within the slug of cells expressing the pDd63-Tag-cat fusion gene. A cloned transformant, with a copy number for the fusion gene of ~100, was developed on agar under conditions which promote slug migration²⁷. After three days of migration the distribution of cells expressing the fusion gene was determined in whole mounts. a and b, Two representative slugs. The staining region is smaller than the prestalk zone as defined by neutral red staining. An analysis of 11 slugs shows the neutral red stained zone occupies, on average, 22% of the length of the slug. After photography, the slugs were individually fixed and stained and the region expressing the pDd63 fusion gene occupied an average of only 13% of slug length.

METHODS. The agar block surrounding an individual slug was excised and fixed in acetone for 2 min. The entire slug was then carefully transferred into a drop of PBS on a poly-L-lysine-coated slide. The buffer was then aspirated and the slide allowed to dry for at least 2 h before immunostaining, as described in the legend to Fig. 1c.

prestalk cell mass at the back of the slug, do not express the pDd63 gene (Fig. 2a, b.)

Staining of whole mounts of slugs, derived from pDd56 transformants, reveals a central 'funnel' of staining cells in the anterior zone, which does not normally extend to the extreme tip and which occupies 10-20% of the length of the slug (Fig. 3a, b). Often, the funnel of cells bends away from the centre (Fig. 3b) and we believe that the cells near the periphery of the slug are being deposited downwards onto the slime trail. The rearguard cells are not stained but we detect a few stained nuclei, scattered throughout the prespore zone, which we assume to be anterior-like cells expressing the pDd56 gene.

Cells continue to accumulate the fusion proteins during extended periods of slug migration of up to three days (Fig. 1b for pDd63; data not shown for pDd56). The prestalk A and B (pstA and pstB) cells retain their characteristic distribution, suggesting that these cells are stable, localized subclasses of the prestalk cell population. This conclusion is based on data which show that the fusion protein itself is stable (K. A. J., unpublished data). PstA cells express the pDd63 gene, but not the pDd56 gene. PstB cells express the pDd56 gene and may also express the pDd63 gene. Because the pstA cells surround the pstB cells, and considerably outnumber them, we have not as yet been able to determine whether the pstB cells express both markers. We have termed the neutral-red stained cells, in the region immediately posterior to the front 10%, pst0 cells (Fig. 3c). These cells express neither marker.

Previous work has suggested the possible existence of multiple classes of prestalk cell, but this is the first study to identify definitive cell-autonomous markers and to localize the cell types. An antibody directed against prestalk-cathepsin, a developmentally regulated cysteine proteinase, binds preferentially to cells in the front half of the prestalk zone¹⁷. The cognate

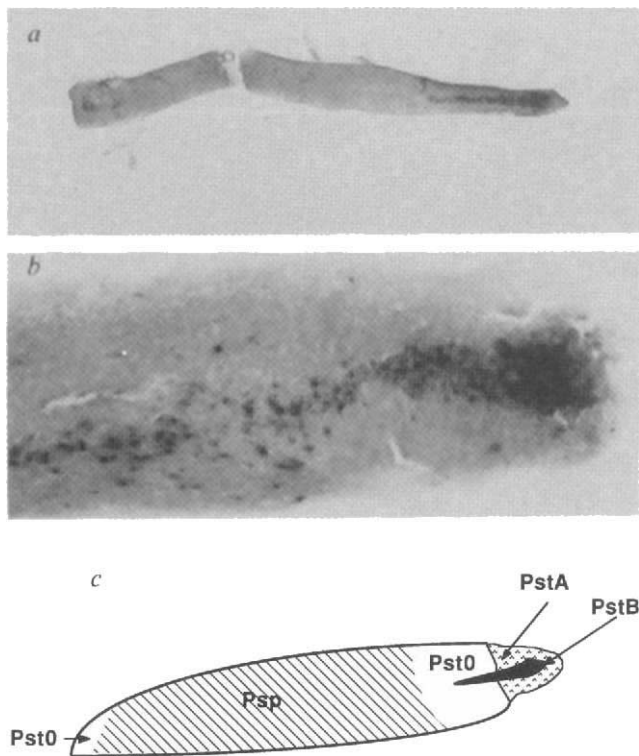


FIG. 3 Analysis of the distribution within the slug of cells expressing the *pDd56-Tag-cat* fusion gene. The cloned transformant described in Fig. 1c was analysed as described in the legend to Fig. 2. a, A representative whole mount slug; b, the anterior regions of a representative whole mount slug; c, a diagrammatic representation of a slug showing cells expressing: *pDd63* (pstA), *pDd56* (pstB), neither marker (pst0) and prespore (psp) cells.

messenger RNA, however, is only slightly more enriched in prestalk than in prespore cells¹⁸. Presumably, the apparent higher enrichment of the protein represents, there, a consequence secondary to cellular differentiation, such as a change in its stability. When purified prestalk cells are incubated *in vitro* at low cell-density with DIF, the stalk-cell inducer, ~90% form stalk cells in the presence of cyclic AMP, but only 20% form stalk cells in its absence¹⁹. This has been adduced as evidence for the existence of cAMP-dependent and cAMP-independent prestalk cell populations.

Our observation of multiple classes of prestalk cells, which are spatially separated and apparently respond to different morphogenetic signals, shows *Dictyostelium* pattern formation to be more complex than previously believed. This observation also has important consequences for understanding how diffusible morphogens might interact to induce cellular differentiation. □

Received 10 April; accepted 12 June 1989.

- Jermyn, K. A., Berks, M., Kay, R. R. & Williams, J. G. *Development* **100**, 745-755 (1987).
- Williams, J. G. *et al. Cell* **49**, 185-192 (1987).
- McRobbie, S. J., Jermyn, K. A., Duffy, K., Blight, K. & Williams, J. G. *Development* **104**, 275-284 (1988).
- McRobbie, S. J., Tilly, R., Blight, K., Ceccarelli, A. & Williams, J. G. *Dev. Biol.* **125**, 59-63 (1988).
- Town, C. D., Gross, J. D. & Kay, R. R. *Nature* **262**, 717-719 (1976).
- Kay, R. R. & Jermyn, K. A. *Nature* **303**, 242-244 (1983).
- Gross, J. D. *et al. Phil. Trans. R. Soc. B295*, 497-508 (1981).
- Morris, H. R., Taylor, G. W., Masento, M. S., Jermyn, K. A. & Kay, R. R. *Nature* **328**, 811-814 (1987).
- Ceccarelli, A. *et al. Nucleic Acids Res.* **15**, 7463-7476 (1987).
- Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. *Cell* **39**, 499-509 (1984).
- Harlow, E., Crawford, L. V., Pim, D. C. & Williamson, N. M. *J. Virol.* **39**, 861-869 (1981).
- Mole, S. E., Gannon, J. V., Ford, M. J. & Lane, D. P. *Phil. Trans. R. Soc. B317*, 455-469 (1987).
- Bonner, J. T. *Am. Nat.* **86**, 79-89 (1952).
- Sternfeld, J. & David, C. N. *Dev. Biol.* **93**, 111-118 (1982).
- Sternfeld, J. & David, C. N. *Differentiation* **20**, 10-21 (1981).
- Devine, K. M. & Loomis, W. F. *Dev. Biol.* **107**, 364-372 (1985).
- Gomer, R. H., Datta, S. & Firtel, R. A. *J. Cell Biol.* **103**, 1999-2015 (1986).

- Pears, C. J., Mahbubani, H. & Williams, J. G. *Nucleic Acids Res.* **13**, 8853-8866 (1985).
- Kwong, L., Sobolewski, A., Atkinson, L. & Weeks, G. *Development* **104**, 121-128 (1988).
- Berks, M. & Kay, R. R. *Dev. Biol.* **126**, 108-114 (1988).
- Norris, K., Norris, F., Christianson, L. & Fill, N. *Nucleic Acids Res.* **11**, 5103-5112 (1983).
- Pears, C. J. thesis, Univ. London (1987).
- Twigg, D. & Sheratt, A. J. *Nature* **283**, 216-218 (1980).
- Knecht, D. A., Cohen, S. M., Loomis, W. F. & Lodish, H. F. *Molec. cell. Biol.* **6**, 3973-3983 (1986).
- Early, A. & Williams, J. G. *Gene* **59**, 99-106 (1987).
- Nellen, W., Silan, C. & Firtel, R. A. *Molec. cell. Biol.* **4**, 2890-2898 (1984).
- Kopachik, W., Dhokia, B. & Kay, R. R. *Differentiation* **28**, 209-216 (1985).
- Gorman, C., Moffat, L. & Howard, B. *Molec. cell. Biol.* **1**, 281-288 (1982).
- Sussman, M. & Lovgren, N. *Expl. Cell Res.* **38**, 97-105 (1965).

ACKNOWLEDGEMENTS. We thank David Lane for advice, Sarah Mole for providing the SV40 constructs, Rob Kay for the chemically synthesized DIF-1 and Mei Wang for permission to include the DIF induction experiment. We also thank Rob Kay, David Lane and Julian Gross for comments on the manuscript.

Acetylcholine analogue stimulates DNA synthesis in brain-derived cells via specific muscarinic receptor subtypes

Avi Ashkenazi, J. Ramachandran* & Daniel J. Capon

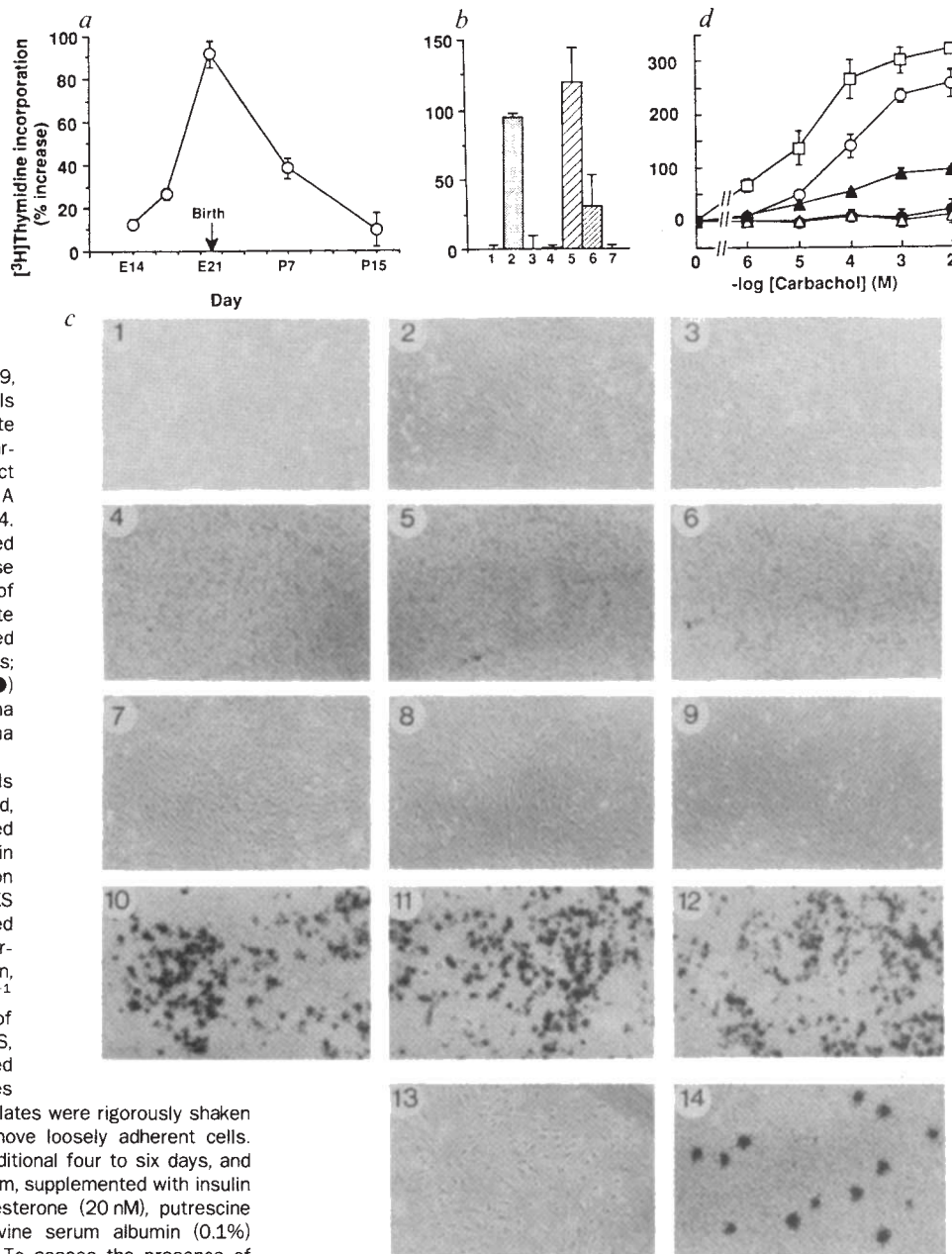
Department of Molecular Biology, Genentech, Incorporated, 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

LITTLE is known about the factors which regulate the growth and development of the mammalian brain. Although proliferation of neuronal cells ceases relatively early in development, certain types of glial cells proliferate and differentiate mainly perinatally¹. In the perinatal period, the ability of acetylcholine to stimulate phosphoinositide (PI) hydrolysis in brain reaches peak levels², and indeed the stable acetylcholine analogue carbachol can stimulate PI hydrolysis of primary neonatal astroglial cells³. As PI hydrolysis is thought to be important in the regulation of cell proliferation⁴⁻⁶, we investigated whether cellular DNA synthesis can be induced by carbachol. Our results show that carbachol stimulates DNA synthesis via muscarinic acetylcholine receptors (mAChRs), in primary astrocytes derived from perinatal rat brain, in an age-dependent fashion. Carbachol is also mitogenic in certain brain-derived astrocytoma and neuroblastoma cell lines, as well as in chinese hamster ovary (CHO) cells expressing recombinant muscarinic receptors. DNA synthesis is strongly activated by carbachol in those brain-derived cell lines and transfected CHO cells that express mAChR subtypes which activate PI hydrolysis efficiently, and poorly activated in cells expressing mAChR subtypes which only weakly activate PI hydrolysis. These results strongly support a role for acetylcholine in regulating astroglial cell growth in the developing brain, and indicate that the specificity of acetylcholine-induced cell proliferation may be determined by the expression of those mAChR subtypes which activate PI hydrolysis.

We examined the ability of carbachol to stimulate DNA synthesis in primary cultures of glial cells derived from the brain of rats between embryonic day (E) 14 and postnatal day (P) 15. The cultures were grown to confluency in the presence of 10% fetal calf serum (FCS), and transferred to a chemically defined serum-free medium⁷ for 48 h. Over 90% of the cells in such cultures are astrocytes, as they stain positively with monoclonal antibodies to the glial fibrillary acidic protein (GFAP)⁸ (data not shown). Carbachol elicited a significant increase in DNA synthesis, observed by increased incorporation of [³H]thymidine, which was detectable in astrocyte cultures from E14 rats, peaked at birth, and returned to baseline at P15 (Fig. 1a). The maximal stimulation by carbachol at birth was ~75% of the stimulation by FCS (Fig. 1b). The effect was mediated by

* Present address: Neurex Corporation, 3760 Haven Avenue, Menlo Park, California, USA.

FIG. 1 Effect of carbachol on DNA synthesis in primary astrocyte cultures and in brain-derived cell-lines-cell. *a*, Effect of carbachol (1 mM) on [3 H]thymidine incorporation in primary cultures of astrocytes derived from the brain of rats from embryonic day (E) 14 to post-natal day (P) 15. *b*, Effects of various agents on [3 H]thymidine incorporation in primary astrocyte cultures from the brain of E21 (newborn) rats: (1) control; (2) carbachol (1 mM); (3) carbachol (1 mM) plus atropine (10 μ M); (4) oxotremorine (1 mM); (5) 10% FCS; (6) noradrenaline (100 μ M); (7) angiotensin II (10 μ M). *c*, Photomicrographs (panels 1–3, 7–9, 13) and corresponding autoradiographs (panels 4–6, 10–12, 14) of single-cell-derived astrocyte colonies from E21 rat brain treated with carbachol (1 mM), showing the variation in the effect of carbachol on [3 H]thymidine incorporation. A higher magnification is shown in panels 13, 14. No response was seen in colonies not treated with carbachol (data not shown). *d*, Dose response analysis of carbachol induction of [3 H]thymidine incorporation in primary astrocyte cultures from E21 rats ($\blacktriangle$) and in brain-derived cell lines: ($\circ$) 1321N1 human astrocytoma cells; ($\square$) SK-N-SH human neuroblastoma cells; ($\bullet$) NG108-15 mouse neuroblastoma X rat glioma hybrid cells; ($\triangle$) N1E-115 mouse neuroblastoma cells.



METHODS. Primary cultures of astroglial cells were prepared as follows: rats were decapitated, the intact brain tissue was removed, separated from the meningeal membranes, and placed in ice cold Hank's physiological salt solution (Sigma) supplemented with 25 mM HEPES pH 7.4. Brain tissue from 3–8 rats was pooled and cut into small pieces. The cells were dispersed by trituration in 0.15% trypsin for 15 min, washed in DMEM containing 4.5 mg ml⁻¹ glucose, 10% FCS (Gibco), 100 U ml⁻¹ each of penicillin and streptomycin and 25 mM HEPES, pH 7.4, filtered to remove aggregates, and plated at high density in 24-well tissue culture dishes precoated with poly-L-lysine. After 24 h, the plates were rigorously shaken for 1 min and washed with medium, to remove loosely adherent cells. Cultures were grown to confluency for an additional four to six days, and transferred to the same medium without serum, supplemented with insulin (5 μ g ml⁻¹), transferrin (100 μ g ml⁻¹), progesterone (20 nM), putrescine (100 μ M), sodium selenite (30 nM) and bovine serum albumin (0.1%) (modified from ref. 7) for a period of 48 h. To assess the presence of astrocytes in the cultures, cells were washed with PBS, fixed with 100% methanol, and stained with an anti-GFAP monoclonal antibody, followed by fluorescein-conjugated rabbit anti-mouse IgG antibodies (Boehringer), and visualized by fluorescence microscopy as previously described²⁸. Over 90% of the cells were GFAP⁺ before or after serum deprivation, indicating that most of the cells in these cultures are astrocytes. Morphological examination (see ref. 1 for review) using phase-contrast microscopy, revealed that after several days of growth in the presence of 10% FCS, 60–80% of the cells formed monolayers and possessed a flat morphology characteristic of type-1 astrocytes, and 20–40% grew on top of these monolayers and possessed a fibrous morphology characteristic of type-2 astrocytes. Bipolar O-2A cells and oligodendrocytes were also seen, but rarely. Under serum-free conditions, the fibrous morphology of some of type-2 cells became more apparent, but the general distribution of cell types did not change significantly. Over 95% of the cells survived at least 72 h in serum-free medium, as determined by trypan blue exclusion. Carbachol and/or other agents were added to the serum-deprived cells, together with [3 H]thymidine (5 Ci mmol⁻¹, 2 μ Ci ml⁻¹), and incubation was continued for additional 24 h. Cells were washed, lysed, and [3 H]thymidine incorporation was measured by trichloroacetic acid precipitation and scintillation counting. Control cells incorporated ~45,000 c.p.m. per well per 24 h. In *c*, confluent cells were detached and transferred to suspension by incubation in PBS containing 5 mM EDTA (15 min, 37 °C) counted, diluted, and re-plated at a low enough density to allow growth of discrete colonies from single cells. After 7 days in culture under the

conditions described above, the colonies were deprived of serum for 48 h, carbachol was added for 24 h, and [3 H]thymidine was included for the last 4 h of carbachol incubation. The cells were washed, methanol-fixed, air dried, and directly autoradiographed (72 h) using Hyperfilm- 3 H (Amersham). A similar experiment, using a shorter exposure, confirmed that the incorporation of [3 H]thymidine was restricted to the cell nuclei. Data from representative experiments (*a*, *b*, *d*, mean $\pm$ s.e.m., $n=8$) or microscopic fields of colonies (*c*) are depicted. SK-N-SH cells were grown in RPMI 1640 medium with 10% FCS, 1321-N1 cells in DMEM with 5% FCS, and N1E-115 and NG108-15 cells were grown in DMEM with 10% FCS, 0.1 mM hypoxanthine, 0.4 μ M bioprotein and 16 μ M thymidine. The media for all cell lines contained 4.5 mg ml⁻¹ glucose, 100 U ml⁻¹ each of penicillin and streptomycin, and 25 mM NaHCO₃ pH 7.4. Confluent cultures were incubated in serum-free medium supplemented with 100 μ g ml⁻¹ transferrin and 0.1% BSA for 48 h, treated with carbachol for another 20 h, [3 H]thymidine (2 μ Ci ml⁻¹) was added for an additional 4 h, and incorporation determined as described above. Data from representative experiments are depicted (means $\pm$ s.e.m., $n=4$). Control cells incorporated from 7,500 to 10,000 c.p.m. per well per 4 h. Carbachol elicited less than 5% increase in [3 H]thymidine incorporation in SK-N-SH and 1321-N1 cells in the presence of 10 μ M atropine; FCS (10%) elicited a 5–7 fold increase in all four cell lines, which was not affected by atropine (data not shown).

TABLE 1 Effects of carbachol on DNA synthesis, PI hydrolysis and adenylyl cyclase activity in cells expressing endogenous or transfected mAChR subtypes

Cell type	mAChR subtype	Number of mAChRs per cell ($\times 10^{-3}$)	DNA synthesis		PI hydrolysis		Adenylyl cyclase	
			Maximal activation (fold)	EC ₅₀ (μ M)	Maximal activation (fold)	EC ₅₀ (μ M)	Maximal inhibition (%)	EC ₅₀ (μ M)
Primary astrocytes	ND	≥ 1	2.0 ± 0.1	30	3.3*	22*	ND	ND
SK-N-SH	M4, M2	8†	4.2 ± 0.1	20	3.0	30	46†	0.01†
1321N1	M4	8‡	3.8 ± 0.1	40	6.2‡	40‡	0‡	—
NG108-15	M3	2†	1.0 ± 0.1	—	1.0†	—	35‡	5.0‡
N1E-115	M3	2†	1.0 ± 0.1	—	1.3†	—	40†	0.40†
CHO	ND	< 0.4	1.0 ± 0.2	—	1.0 ± 0.2	—	0	—
CHO	M1	16	1.7 ± 0.1	200	ND	ND	0	—
CHO	M1	26	2.3 ± 0.1	60	5.4 ± 0.3	30	0	—
CHO	M1	50	4.9 ± 0.1	100	13.0 ± 0.7	40	0	—
CHO	M2	16	1.0 ± 0.1	—	1.1 ± 0.1	—	76	0.03
CHO	M2	75	1.6 ± 0.1	100	2.0 ± 0.1	30	70	0.01
CHO	M2	1,000	2.5 ± 0.3	40	3.7 ± 0.1	40	72	0.07
CHO	M3	14	1.0 ± 0.2	—	1.3 ± 0.1	—	48	0.07
CHO	M3	80	1.5 ± 0.1	10	2.1 ± 0.3	5	69	0.01
CHO	M4	26	1.9 ± 0.1	100	7.5 ± 2.0	10	0	—
CHO	M4	190	5.5 ± 0.3	20	18.5 ± 0.6	10	0	—

The levels of mAChRs were measured by saturation binding analysis of the nonselective mAChR antagonist [³H]N-methylscopolamine (NMS) as described^{13,17}. Cells were incubated with various concentrations of [³H]NMS for 90 min at 37 °C, washed with phosphate buffered saline (PBS, pH 7.4), pelleted, and dissolved in Aquasol-2 (NEN) for determination of radioactivity by scintillation counting. Non-specific binding was determined in the presence of 10 μ M atropine. The level of [³H]NMS binding to primary astrocytes was not significantly altered by the cell density, within a range of 2×10^3 to 1×10^5 cells per cm² (data not shown). Activation of DNA synthesis was measured as described in Fig. 1; the fold activation is based on the ratio of [³H]thymidine incorporation in cells incubated with or without carbachol (10 mM). Control levels of incorporation in the various cell types ranged from 7,500 to 13,500 c.p.m. per well per 4 h, each well of a 24 well dish containing $2-3 \times 10^5$ cells. PI hydrolysis was determined by measuring the accumulation of [³H]inositol mono-, bis and triphosphates in cells labelled with [³H]myo-inositol (2 μ Ci/ml⁻¹, 24 hours), over a 30 min incubation with carbachol at 37 °C in the presence of 10 mM LiCl, as described previously^{18,36}. The fold activation is the ratio of total inositol phosphates accumulation in cells treated versus not treated with carbachol (1 mM). Basal levels of inositol phosphate accumulation in the various cell types ranged from 9,000 to 12,000 c.p.m. per 10⁶ cells per 30 min. Inhibition of adenylyl cyclase was determined by measuring the effect of carbachol on cAMP accumulation induced by forskolin (10 μ M) in intact cells treated with the phosphodiesterase inhibitor isobutylmethylxanthine (IBMX) (100 μ M); cAMP was measured by radioimmunoassay as described previously¹⁷; basal cAMP levels were 5–10 pmol per 10⁶ cells; forskolin stimulation of these levels ranged from 9- to 13.5-fold. The fold inhibition is based on the ratio of cAMP levels in cells treated with IBMX, forskolin and carbachol (1 mM) versus cells treated with IBMX and forskolin. Carbachol alone had no detectable effect on cAMP levels. The data from the present investigation are means ($\pm$ s.d. where given) from two to five experiments. ND, not determined. The designation of human mAChR subtypes used here is based on our original report on the cloning and sequencing of four human mAChR subtype genes¹³. The subtypes that we refer to as M3 and M4, have been referred to as m4 and m3, respectively, by Bonner *et al.* (ref. 14), and mAChR IV and mAChR III, respectively, by Neher *et al.* (ref. 21).

* Data are from ref. 3; † Data are from J. Winslow *et al.*, manuscript in preparation; ‡ Data are from ref. 34; § Data are from ref. 35.

mAChRs, as it was abolished by the muscarinic antagonist atropine (Fig. 1b). Specific binding of the mAChR ligand [³H]N-methylscopolamine indicated that these primary astrocyte cultures express $\sim 1,000$ mAChRs per cell, confirming significant mAChR expression in at least some astrocytes (Table 1). Oxotremorine, a muscarinic agonist, which unlike carbachol does not stimulate PI hydrolysis, did not stimulate DNA synthesis; noradrenaline stimulated DNA synthesis by $\sim 30\%$, whereas angiotensin II had no effect (Fig. 1b). Autoradiography of individual astrocyte colonies grown from single cells from neonatal brain revealed that some colonies respond to carbachol well, and others poorly (Fig. 1c).

One possible reason for the lack of response in some of the astrocyte colonies could be that they lack mAChRs. We have found however, that most glial, as well as neuronal, cell lines express mAChRs (ref. 9 and J. Winslow *et al.*, unpublished results). Because the mAChR family is now known to consist of at least five receptor subtypes, with distinct primary structures^{10–15}, ligand-binding and effector-coupling properties^{13,15–24}, and patterns of tissue-specific expression^{13–15,25}, we reasoned that differences in mAChR subtype expression by individual cells might contribute to the variation in response. We therefore took advantage of our recent finding that brain-derived glial and neuronal cell lines show distinct patterns of mAChR subtype messenger RNA expression^{9,13}, to investigate the ability of carbachol to stimulate DNA synthesis in cell lines expressing either the M3 subtype, the M4 subtype or both the M4 and M2 subtypes. 1321N1 astrocytoma cells, and SK-N-SH neuroblastoma cells, which express M4, responded to carbachol

with $\sim 300\%$ increase in DNA synthesis, whereas NG108-15 neuroblastoma X glioma and N1E-115 neuroblastoma cells, which express only M3, did not respond (Fig. 1d, Table 1). Carbachol stimulation was blocked by atropine (Fig. 1 legend). The differential response of these brain cell lines to carbachol appeared to be due to the presence or absence of the M4 subtype, but could also be a coincidence, and due to some other difference between the cell types.

To evaluate these possibilities, we investigated whether the expression of different mAChRs in a single cell type confers differential sensitivity to mitogenic stimulation by carbachol. We transfected CHO cells, which lack endogenous mAChRs^{12,17}, with the human M1, M2, M3 or M4 mAChR genes¹³, to establish cell lines stably expressing individual mAChR subtypes (Table 1). DNA synthesis was stimulated by carbachol in CHO cells expressing each of the four mAChR subtypes, but not in untransfected cells; however, the response was much greater in cells expressing M1 or M4 than in cells expressing M2 or M3 (Fig. 2). For cells expressing each mAChR subtype, the response increased with receptor number, but for M2 or M3, the number of receptors required to give an appreciable response was markedly higher than for M1 and M4 (Table 1). Thus, mAChR subtypes differ significantly in their ability to activate DNA synthesis in a given cell, providing strong support for the notion that the differential mitogenic response seen in brain cell lines results from expression of specific mAChR subtypes. Therefore, DNA synthesis in 1321N1 and SK-N-SH cells is strongly activated by carbachol probably as a consequence of the expression in these cells of the M4 mAChR, which mediates the mitogenic

TABLE 2 Effect of pertussis toxin (PTX) on carbachol stimulation of DNA synthesis

	³ H]thymidine incorporation					
	M1		M2		M4	
	(c.p.m. per well)	(fold)	(c.p.m. per well)	(fold)	(c.p.m. per well)	(fold)
Control	13,360 ± 310	1.0	9,970 ± 440	1.0	10,500 ± 640	1.0
PTX	14,760 ± 170	1.1	10,210 ± 160	1.1	10,840 ± 560	1.0
Carbachol	52,390 ± 1,290	3.9	20,670 ± 1,660	2.1	55,940 ± 1,300	5.3
Carbachol + PTX	45,680 ± 660	3.4	10,040 ± 130	1.0	35,120 ± 1,440	3.3
Carbachol + atropine	14,700 ± 1,260	1.1	9,600 ± 360	1.1	10,940 ± 260	1.0

The effect of PTX was analysed in CHO cell lines expressing the M1, M2 or M4 mAChRs (5×10^4 , 1×10^6 and 1.9×10^5 receptors per cell, respectively). DNA synthesis as assayed as in Fig. 2. Cells were treated with PTX (100 ng per ml) for 6 h before the addition of carbachol. This treatment results in complete ADP-ribosylation of any detectable PTX-sensitive G protein in these cells¹⁷. The concentrations in the assay were 1 mM for carbachol and 10 μ M for atropine. Data from representative experiments (means $\pm$ s.e.m., $n=4$) are depicted.

signal efficiently. NG108-15 and N1E-115 cells do not respond however, because they express the relatively inefficient M3 mAChR subtype.

To determine whether this difference between the mAChR subtypes is related to differences in their effector-coupling properties, we compared the signalling pathways activated by each mAChR subtype in CHO cells. Recent studies with transfected cells indicate that the M1 and M4 subtypes are substantially better than M2 and M3 in activating PI hydrolysis¹⁶⁻²⁴, a signal transduction pathway which has been implicated in the induction of cell proliferation⁴⁻⁶. PI hydrolysis was efficiently activated by carbachol in M1- and M4-expressing CHO cells at relatively low receptor numbers, but only weakly activated in CHO cells expressing even high levels of M2 and M3 receptors; adenylyl cyclase was inhibited efficiently by M2 and M3, but not by M1 and M4 mAChRs (Table 1). Thus, the degree to which DNA synthesis was stimulated by carbachol correlated well with the degree of activation of PI hydrolysis, and was not related to adenylyl cyclase inhibition. Furthermore, the carbachol EC_{50} s for PI hydrolysis and DNA synthesis were similar, whereas inhibition of adenylyl cyclase occurred at much lower carbachol concentrations (Fig. 1d, Table 1). Both the magnitude and carbachol EC_{50} of PI hydrolysis and DNA synthesis also correlated well in the brain cell lines (Table 1), as well as in the neonatal primary astrocyte cultures (carbachol EC_{50} for stimulation of DNA synthesis is 30 μ M (Fig. 1d), and for PI hydrolysis is 22 μ M (ref. 3)). These results strongly suggest that mAChR activation of PI hydrolysis, rather than inhibition of adenylyl cyclase, is involved in the mitogenic action of carbachol. The observation that oxotremorine has no mitogenic activity in primary astrocyte cultures (Fig. 1b) supports this interpretation, as oxotremorine is known to induce mAChR-mediated inhibition of adenylyl cyclase, but not activation of PI hydrolysis, in brain and heart tissue^{26,27}.

If PI hydrolysis is involved in the activation of DNA synthesis by carbachol, its uncoupling from agonist binding to the mAChR should result in a loss of the mitogenic response. To test this prediction, we examined the effect of pertussis toxin (PTX), which selectively uncouples different mAChR subtypes from PI hydrolysis in CHO cells²⁴, on carbachol stimulation of DNA synthesis. Indeed, in CHO cells expressing 1×10^6 M2 receptors per cell, in which carbachol stimulation of PI hydrolysis is completely abolished by PTX²⁴, carbachol-stimulated DNA synthesis was also completely abolished (Table 2). In cells expressing 50,000 M1 or 190,000 M4 receptors per cell, in which ~20% or 50%, respectively, of carbachol-stimulated PI hydrolysis is inhibited by PTX²⁴, a proportional PTX inhibition of carbachol-stimulated DNA synthesis was seen (20% or 45%, respectively) (Table 2). Thus, the degree to which PI hydrolysis and DNA synthesis are stimulated by carbachol is proportional, the carbachol EC_{50} s for these responses are similar, and PTX inhibits carbachol stimulation of both responses to a comparable extent.

Together, these findings strongly support the conclusion that PI hydrolysis is involved in the mitogenic stimulation by carbachol.

Our demonstration that carbachol stimulates DNA synthesis in primary neonatal astrocyte cultures provides evidence that acetylcholine is important in astroglial cell growth during brain development. The stimulation by carbachol is most apparent in astrocyte cultures from newborn rat brain, which temporally coincides with the occurrence of specific types of astrocytes in the brain during late embryogenesis and the neonatal period (reviewed in ref. 1). Type-1 (or protoplasmic) astrocytes first appear in the brain at E15-16, and divide frequently for at least one week, whereas type-2 (or fibrous) astrocytes develop from a different progenitor cell, first appear at P7-10 and divide infrequently^{1,28}. The other major type of glial cells, the oligodendrocytes, develop from the same type of progenitor (O-2A cells) as type-2 astrocytes, first appear at birth, and do not proliferate^{1,28}. In cultures of E ≥ 16 brain, grown in the presence of 10% FCS, type-1 astrocytes develop on the same schedule

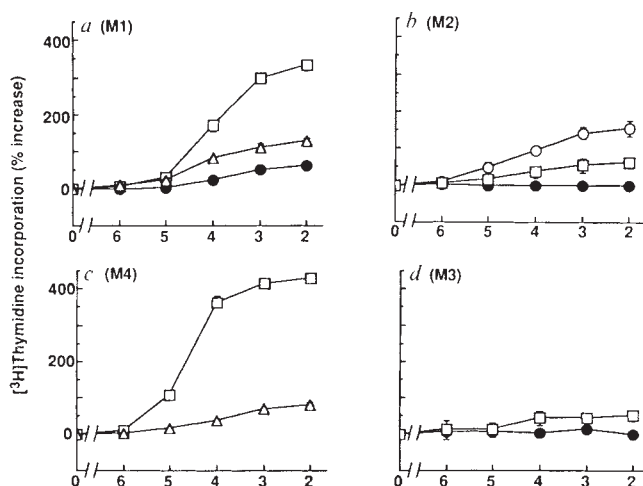


FIG. 2 Effect of carbachol on DNA synthesis in CHO cells transfected with individual muscarinic receptor subtypes. a, Cell lines expressing 16,000 (●), 26,000 (△) or 50,000 (□) M1 receptors per cell; b, cell lines expressing 16,000 (●), 75,000 (□) or 1,000,000 (○) M2 receptors per cell; c, cell lines expressing 26,000 (△) or 190,000 (□) M4 receptors per cell; d, cell lines expressing 14,000 (●) or 80,000 (□) M3 receptors per cell. Data from representative experiments are depicted (means $\pm$ s.e.m., $n=4$). METHODS. CHO cells were grown, transfected, and clonal mAChR-expressing lines were selected as previously described^{12,17}. Serum-free medium contained human transferrin (100 μ g ml⁻¹), BSA (0.1%) and HEPES buffer (25 mM, pH 7.4). Control cells incorporated from 10,000 to 13,500 c.p.m. per well per 4 h.

as *in vivo*, whereas O-2A cells differentiate prematurely (within 1–2 days) into type-2 astrocytes, and oligodendrocytes rarely develop²⁸. Thus, the conditions that we have used (5–7 days of culture in the presence of 10% FCS) favour the development of type-1 and type-2 astrocytes. Indeed, both cell types could be identified in our cultures based on their distinct morphology¹. Type-1 cells were more abundant than type-2 cells, whereas very few O-2A cells or oligodendrocytes were observed (Fig. 1 legend). Because type-1 astrocytes continue to proliferate for at least a week after they develop, whereas type-2 astrocytes do not¹, it is most likely that those cells which respond to carbachol *in vitro* are type-1 astrocytes (or their progenitors). This possibility is strongly supported by the observation that carbachol can stimulate DNA synthesis in colonies derived from single cells having the characteristic flat morphology of type-1 astrocytes (see Fig. 1c, panels 13, 14). This also indicates that the effect of carbachol on these cells is direct. Although it is unlikely that the effect on DNA synthesis seen here is due to stimulation of O-2A cells, as most of these cells differentiate rapidly in the presence of 10% FCS, it is nonetheless possible that these cells are affected by acetylcholine *in vivo*, either directly, or indirectly through type-1 astrocytes. Type-1 astrocytes are known to secrete platelet-derived growth factor (PDGF), which induces O-2A cells to proliferate^{1,28}; therefore an increased number of type-1 astrocytes *in vivo* may elevate the levels of PDGF, stimulating O-2A cell proliferation.

Our results indicate that the mitogenic action of carbachol depends not only on mAChR expression, but also on the particular subtypes of mAChR which are expressed in the individual cell. Consistent with the conclusion that PI hydrolysis is important in the regulation of astrocyte proliferation, we have also found significant stimulation of DNA synthesis in primary astrocytes by noradrenaline, which stimulates PI hydrolysis via α -1 adrenergic receptors in these cells³. Although noradrenaline was less potent than carbachol, it may act as a mitogen at other stages of development, or in other types of brain cells, as may other G-protein-linked neurotransmitters that regulate PI hydrolysis. This possibility is supported by the previous observations that serotonin²⁹ and bradykinin³⁰ are mitogenic in fibroblasts, and that the *mas* oncogene encodes an angiotensin receptor³¹.

The regulation of astroglial cell proliferation by acetylcholine provides a novel example of communication between neurons and glial cells in the central nervous system. Among the important, neural-supportive functions described to date for astrocytes is their involvement in the uptake and metabolism of neurotransmitters such as glutamate and γ -aminobutyric acid (reviewed in ref. 32). Astrocytes were known to possess α - and β -adrenergic, muscarinic, and other neurotransmitter receptors, but the biological consequences of signal transduction by these receptors in astrocytes were not previously known³². Our results indicate that one function of astrocyte mAChRs may be to coordinate the development of specific types of astroglial cells with that of cholinergic neurons in the brain. Consistent with this possibility, acetylcholine is present in significant amounts in the neonatal brain (~30% of adult levels)³³, and presumably capable of exerting local effects on astrocytes. Thus, certain neurotransmitters may act as neuronal-glial communicators in the developing brain, in addition to their roles in communication between neurons and synaptic plasticity in the adult brain. □

11. Kubo, T. *et al.* *FEBS Lett.* **209**, 367–372 (1986).
12. Peralta, E. G. *et al.* *Science* **236**, 600–605 (1987).
13. Peralta, E. G. *et al.* *EMBO J.* **6**, 3923–3929 (1987).
14. Bonner, T. I., Buckley, N. J., Young, A. C. & Brann, M. R. *Science* **237**, 527–532 (1987).
15. Bonner, T. I., Young, A. C., Brann, M. R. & Buckley, N. J. *Neuron* **1**, 403–410 (1988).
16. Fukuda, K. *et al.* *Nature* **327**, 623–625 (1987).
17. Ashkenazi, A. *et al.* *Science* **238**, 672–675 (1987).
18. Peralta, E. G., Ashkenazi, A., Winslow, J. W., Ramachandran, J. & Capon, D. J. *Nature* **334**, 434–438 (1988).
19. Fukuda, K. *et al.* *Nature* **335**, 355–358 (1988).
20. Stein, R., Pinkas-Karamaski, R. & Sokolovsky, M. *EMBO J.* **7**, 3031–3035 (1988).
21. Neher, E., Marty, A., Fukuda, K., Kubo, T. & Numa, S. *FEBS Lett.* **240**, 88–94 (1988).
22. Bujo, H. *et al.* *FEBS Lett.* **240**, 95–100 (1988).
23. Conklin, B. R. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 8698–8702 (1988).
24. Ashkenazi, A., Peralta, E. G., Winslow, J. W., Ramachandran, J. & Capon, D. J. *Cell* **56**, 487–493 (1989).
25. Maeda, A., Kubo, T., Mishina, M. & Numa, S. *FEBS Lett.* **239**, 339–342 (1988).
26. Fisher, S. K., Klinger, P. D. & Agranoff, B. W. *J. biol. Chem.* **258**, 7358–7363 (1983).
27. Brown, J. H. & Brown, S. L. *J. biol. Chem.* **259**, 3777–3781 (1984).
28. Williams, B. P., Abney, E. R. & Raff, M. C. *Dev. Biol.* **112**, 126–134 (1985).
29. Seuwern, K., Magnaldo, I. & Pouyssegur, J. *Nature* **335**, 254–256 (1988).
30. Vincentini, L. M. & Villereal, M. L. *Biochem. biophys. Res. Commun.* **123**, 377–384 (1984).
31. Jackson, T. R., Blair, L. A. C., Marshall, J., Goedert, M. & Hanley, M. R. *Nature* **335**, 437–440 (1988).
32. Kimelberg, H. K. & Norenberg, M. D. *Scient. Am.* **260**, 66–76 (1989).
33. Coyle, J. T. & Yamamura, H. I. *Brain Res.* **118**, 429–440 (1976).
34. Masters, S. B., Harden, T. K. & Brown, J. J. *Molec. Pharmac.* **26**, 149–155 (1984).
35. Lichtshtein, D., Boone, G. & Blume, A. *J. Cyc. Nucleotide Res.* **5**, 367–375 (1979).
36. Berridge, M. J., Downes, C. P. & Hanley, M. R. *Biochem. J.* **206**, 587–595 (1982).

ACKNOWLEDGEMENTS. We thank Scot Marsters for tissue culture assistance, Drs Ernest Peralta and John Winslow for discussions, and Dr Rebecca Ward for reading the manuscript.

Effect on eye development of dominant mutations in *Drosophila* homologue of the EGF receptor

Nicholas E. Baker & Gerald M. Rubin

Howard Hughes Medical Institute and Department of Biochemistry, University of California, Berkeley, California 94720, USA

THE compound eye of the adult fruitfly, *Drosophila melanogaster*, comprises about 800 identical ommatidia, or unit eyes, each containing 20 distinct cells^{1,2}. We have used histological and immunocytochemical methods to study the development of the compound eye in *Ellipse* (*Elp*) mutants. In *Elp/Elp*, most ommatidia do not initiate differentiation. We present genetic evidence that *Elp* alleles are mutations of the *Drosophila* homologue of the epidermal growth factor (EGF) receptor, and suggest that activity of the EGF receptor may determine the spacing pattern of ommatidia in the eye.

In *Drosophila* each eye derives from a sheet of undifferentiated epithelial cells called the eye imaginal disk. Cell determination and differentiation begin towards the end of the third and final larval instar. The first sign of developing ommatidia is the formation of an array of preclusters containing five of the eventual eight photoreceptor cells. Some cells begin to express specific neural antigens at this stage. Subsequently more photoreceptors and other cell types are added to each precluster. Within each ommatidium, cell interactions are thought to determine the specific identity of each cell. How preclusters come to form in a regular spatial array is unknown, but local competition and lateral inhibition between disk cells may be involved^{1,2}.

The two *Ellipse* mutations *Elp*¹ and *Elp*^{B1} both affect compound eye morphology. In *Elp/+* heterozygotes the array of ommatidia is less regular than in the wild type (Fig. 1, compare panels *a* and *b*). There is also a slight disturbance of the wing-vein pattern (data not shown). *Elp/Elp* homozygotes show a much more extreme defect (Fig. 1c). The smaller eyes contain many fewer ommatidia and some regions lack them entirely. Sections of fixed material show that most ommatidia contain the normal number and arrangement of cells. The regions without facets contain cells resembling the pigment cells that usually surround each ommatidium, and also mechanosensory bristles (Fig. 1d–f).

Received 30 January; accepted 25 April 1989.

1. Raff, M. C. *Science* **243**, 1450–1455 (1989).
2. Balduini, W., Murphy, S. D. & Costa, L. G. *J. Pharmac. exp. Ther.* **241**, 421–427 (1987).
3. Pierce, B. *et al.* *J. Neurochem.* **45**, 1534–1539 (1985).
4. Rozengurt, E. *Science* **234**, 161–166 (1986).
5. Berridge, M. J. *A. Rev. Biochem.* **56**, 159–193 (1987).
6. Nishizuka, Y. *Nature* **334**, 661–665 (1988).
7. Bottenstein, J. E. & Sato, G. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 514–517 (1979).
8. Bingami, A., Eng, L. F., Dahl, D. & Uyeda, C. T. *Brain Res.* **43**, 429–435 (1972).
9. Ashkenazi, A., Peralta, E. G., Winslow, J. W., Ramachandran, J. & Capon, D. J. *Cold Spring Harb. Symp. quant. Biol.* **LIII**, 263–272 (1988).
10. Kubo, T. *et al.* *Nature* **323**, 411–416 (1986).

Specific antibodies can be used to determine the pattern of differentiation in eye imaginal disks¹⁻³. Disks from wild-type and *Elp/Elp* larvae were stained with antibodies that label identified cells at the precluster stage^{4,5}. The results are illustrated in Fig. 2. In the *Elp* homozygotes only about one tenth of the normal number of preclusters differentiated, and the remaining eye disk cells were indistinguishable from the undifferentiated cells that separated preclusters in wild-type disks. The few preclusters formed seemed to continue development as normal, because at later stages further photoreceptor cells and cone cells were added. Other cells, however, remained undifferentiated by the start of puparium formation. Some of these undifferentiated cells die, whereas others must give rise to the pigment cells and bristles seen in the adult.

It has been suggested previously that within each ommatidium, cell fate is specified by short-range interactions autonomous to each cluster^{1,4}. The description of *Elp* provides strong support for this view, because in *Elp/Elp* disks ommatidia can develop normally when separated from others (see Figs 1 and 2 for examples).

Deletion chromosomes were used to investigate the effect of gene dosage on the *Elp* mutations. The dominant *Elp* phenotype was reverted when *Df(2R)D17* was induced in an *Elp¹* chromosome, indicating that the *Elp* locus lies within this deficiency

(chromosome bands 57B5-58B1-2)⁶. *DfD17/+* flies had a normal eye, so the dominant *Elp* phenotype is not caused by a reduction of gene dosage. *Elp/DfD17* individuals also had normal eyes (Fig. 1*f*), so the dominant phenotype can be suppressed by lowering the amount of wild-type product. This suggests that the *Elp* phenotype is caused by increased gene activity in the mutant⁷. The suppression of the dominant phenotype by deficiencies was used further to map the *Elp⁺* locus more precisely. Only deletions including the region 57D11-12; 57F5-6 had this property, suggesting the *Elp⁺* locus lies within this interval (we used *Df(2R)D17*, *Df(2R)F36*, *Df(2R)AA21*, *Df(2R)PK1* and *Df(2R)PL3*; ref. 6).

Because *DfD17/+* flies had normal eyes, loss-of-function mutations could be obtained by reverting *Elp* mutations. Five independent revertants of *Elp¹* or *Elp^{B1}*, induced with either ethylmethane sulphonate or X-rays, define a lethal complementation group also uncovered by deficiencies for the 57D11-12; 57F5-6 interval (see legend to Fig. 3 for details). All of these revertants (*Elp^R*) behave like deficiencies for the locus in suppressing the dominant *Elp* phenotype in *Elp/Elp^R* transheterozygotes (for example, 1*i*), and are presumed to eliminate or greatly reduce the amount of wild-type gene activity. *Elp^R/Elp^R* or *Elp^R/Df* embryos died before hatching. The pattern of larval cuticle from dead embryos was abnormal

FIG. 1 Aspects of eye structure in *Ellipse*. *a*, Scanning electron micrograph of wild-type (Canton S) eye, showing regular array of ommatidia. In panels *a-e*, anterior is to the left and dorsal uppermost. *b*, The *Elp^{B1}/+* eye resembles wild-type but can always be distinguished by the irregular disposition of ommatidia. We induced the *Elp^{B1}* mutation by treatment of sperm with 30 mM ethylmethane sulphonate¹⁵. The *Elp¹* mutation arose spontaneously¹⁶. *Elp^{B1}* and *Elp¹* are allelic because of their identical phenotypes and map positions (2-99), and because recessive revertants of both mutants belong to the same complementation group (see text and legend to Fig. 3). *c*, In *Elp^{B1}/Elp^{B1}*, the eye is smaller and contains many fewer facets. Regions without facets often been numerous bristles. *Elp¹/Elp¹* and *Elp¹/Elp^{B1}* are similar. *d*, Section through a wild-type eye. The photoreceptor cells of each ommatidium are separated by a honeycomb-like network of pigment cells. Scale bar, 100 μ m. *e*, In an *Elp^{B1}/Elp^{B1}* eye, large regions of the eye lack ommatidia and are filled with pigmented cells. Most of the ommatidia formed contain the normal number and arrangement of cells (for example, *f*), but in about one third there is some difference from the wild-type pattern. *f*, Detail from an *Elp^{B1}/Elp^{B1}* section showing an isolated ommatidium. The photoreceptor cells have the same morphology and arrangement as in wild-type ommatidia. The cells surrounding the ommatidium resemble normal pigment cells morphologically. *g*, Enlargement ($\times 4$) of the wild-type eye shown in *a*, illustrating the regular array of ommatidia. *h*, Detail of an *Elp¹/+* eye. Like the *Elp^{B1}/+* eye shown in *b*, the facets and bristles are irregularly arranged and sometimes of unusual sizes. *i*, *Elp*/deficiency eyes are indistinguishable from the wild-type eye shown in *a* and *g*. The same is true of transheterozygotes between *Elp* and loss-of-function mutations (see Fig. 3). An *Elp¹/Elp^{RE124}* eye is shown in this picture.

METHODS. For sections, eyes were fixed and embedded for 1 μ m plastic sections as described previously¹⁷, and stained with toluidine blue. For scanning electron microscopy, whole flies were dried by desiccation and then coated using standard methods.

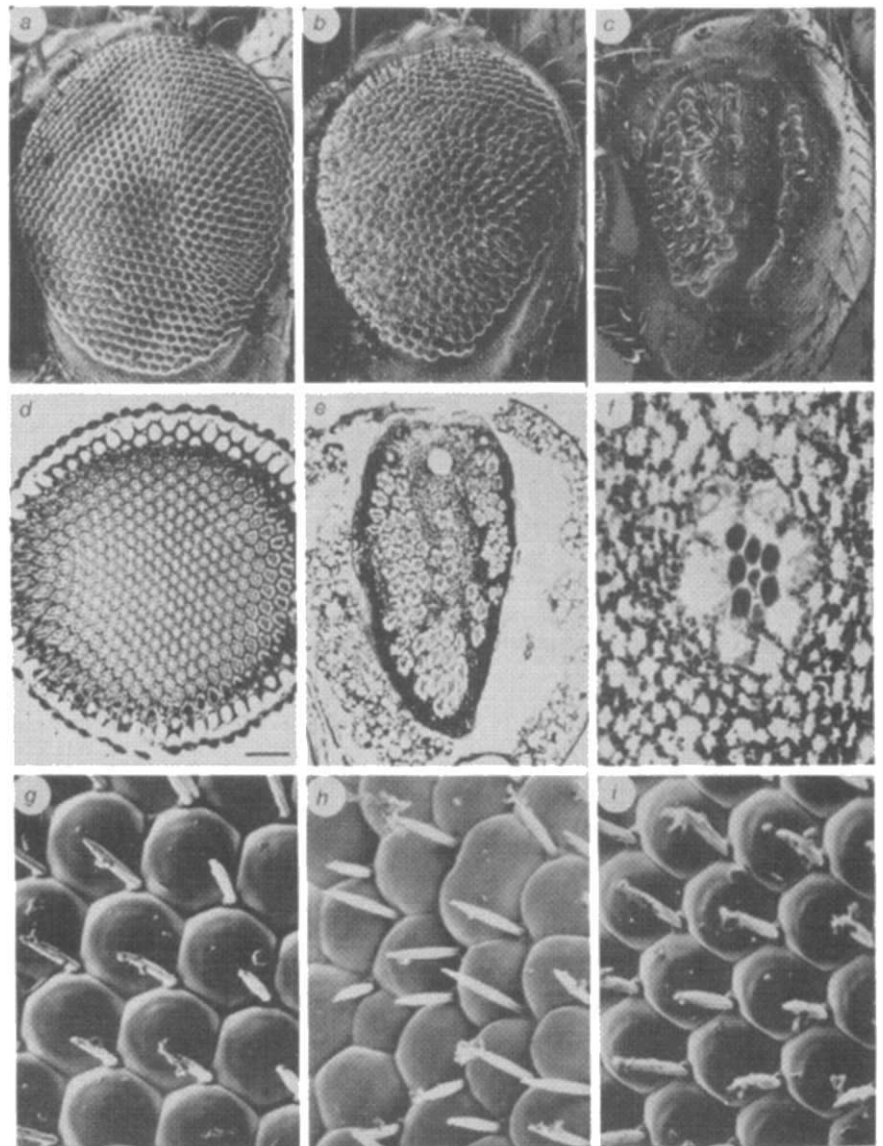
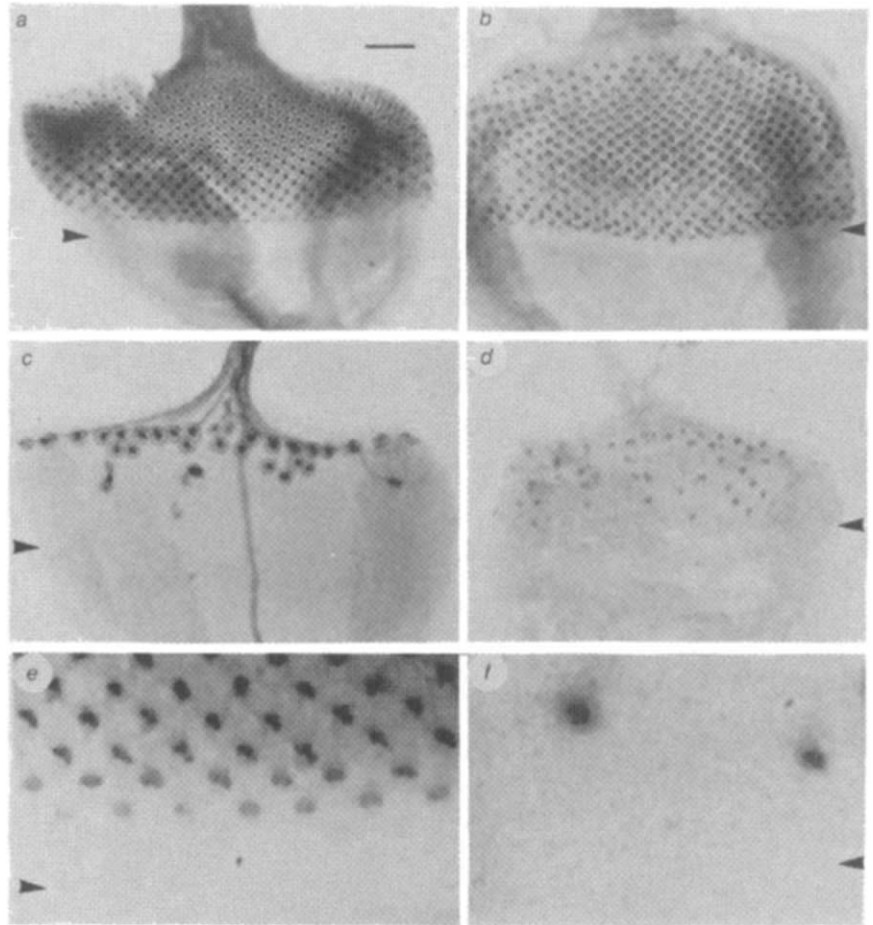


FIG. 2 Antibody staining of imaginal disks. The eye disk begins differentiation in the third larval instar, but all regions of the disk do not develop simultaneously. Instead a groove called the morphogenetic furrow moves from the posterior margin (towards the top of these pictures) to the anterior (towards the bottom), taking about two days to cross the disk. Pattern formation begins as the morphogenetic furrow passes, so that pre-clusters are formed next to the furrow and more advanced stages are formed more posteriorly^{1,2}. *a*, A wild-type (Canton S) eye disk stained with monoclonal antibody BP104. BP104 recognizes a cytoplasmic determinant in all cells of the *Drosophila* peripheral nervous system (A. Bieber, N. Patel and C. S. Goodman, personal communication). Posterior to the morphogenetic furrow (arrowhead), differentiating photoreceptor cells are stained in each ommatidium. Scale bar, 100 μ m for *a-d*. *b*, A wild-type disk stained with an antiserum raised against the C-terminal 13 amino acids of the *sevenless* protein. Posterior to the morphogenetic furrow, the antibody labels several cells within each ommatidium in a pattern described previously⁵. *c*, An *Elp*¹/*Elp*¹ disk stained with BP104. Only scattered ommatidia have differentiated after passage of the morphogenetic furrow (arrowheads). In the ommatidia that do form, more cells are differentiated in the older ommatidia away from the furrow, as in the wild type. In the majority, the arrangement of photoreceptor cells is the same as normal, but sometimes there are differences (not indicated). *d*, Staining with antibody against the *sevenless* protein also shows differentiation of only a few ommatidia in an *Elp*¹/*Elp*¹ disk. *e*, Enlargement of the region close to the morphogenetic furrow in a wild-type disk labelled with BP104, showing the array of preclusters separated by undifferentiated cells. *f*, In a similar region from an *Elp*¹/*Elp*¹ disk, most cells have not differentiated as neurons. This picture shows dispersed ommatidia developing in apparent isolation. Fixation and antibody staining of disks were as described previously⁴.

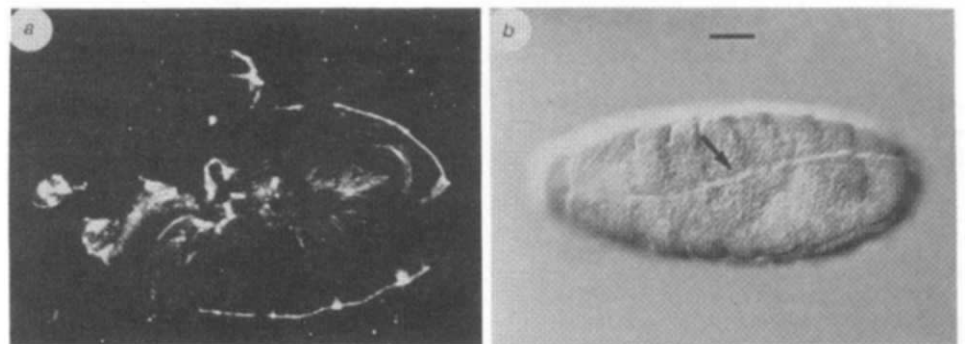


(Fig. 3*a*). The ventral denticle belts were greatly reduced or absent, but fine hairs typical of the dorsal epidermis were present. We could first distinguish *Elp*^R/Df from wild type when the mutant embryos were twisted asymmetrically at the extended germ-band stage (Fig. 3*b*). The phenotype of *Elp*^R/*Elp*^R embryos resembles that of *faint little ball* (*flb*) mutations, which were identified in a previous screen for embryonic lethal mutations affecting pattern formation⁸. The *flb* locus maps to the

same genetic and cytological position as *Elp*. Two alleles of *flb* (*flb*^{2C82}, ref. 8, and *l(2)57DEFa-4*, ref. 6) failed to complement the *Elp*^R mutants, and also suppressed the dominant eye phenotype in *Elp*/*flb* flies, showing that these are all mutations in the same gene.

It has recently been reported that the *flb* locus encodes the *Drosophila* homologue of the EGF receptor (DER)^{9,10}. Our results indicate that both *Elp* alleles are also mutations of the

FIG. 3 Embryonic lethal phenotype of *Elp*^R mutations. *a*, Cuticle pattern produced by an *Elp*^{B1RB5}/*Elp*^{B1RB5} embryo. Anterior to the left, dorsal uppermost. Germ-band retraction has not been completed, so that the abdomen is still folded around the posterior end of the embryo, and the head has not involuted. The ventral denticle bands are greatly reduced. Fine hairs characteristic of dorsal regions are present. *b*, Ventral aspect of an extended germ-band stage embryo, showing distortion of the ventral midline in an *Elp*^{B1RB3}/Df D17 embryo (arrow); anterior to the left. See ref. 18 for descriptions of wild-type embryos. Scale bar, 100 μ m.



METHODS. The *Elp*^{B1} mutation was reverted using 30 mM ethylmethane sulphonate. *Elp*^{B1}/CyO males were mutagenized and mated with Canton S wild-type females. Six independent revertants that no longer had a dominant phenotype were found amongst 9,000 progeny. Four of these (called *Elp*^{B1RB2-5}) were lethal in combination with each other or deficiencies for the 57D11-12; 57F5-6 interval. The *Elp*^{RE124} revertant was obtained by Dr

J. O'Donnell (personal communication) after X-ray irradiation of an *Elp*¹ chromosome (4,000 rad)⁶. We found that *Elp*^{RE124} belongs to the same lethal complementation group. The salivary chromosomes *Elp*^{RE124} and *Elp*^{B1RB3}, and also *flb*^{2C82}, were examined for cytological abnormalities, but none was detected. Embryo fixations and cuticle preparation were performed as described¹⁸.

DER gene. It is important to note that the dominant *Elp* phenotype was suppressed by reducing the amount of wild-type gene product, showing that *Elp* mutations enhanced a function normally active in the same cells. This implies that the *DER* has a normal role in the developing eye. Indeed, *in situ* hybridization has shown the *DER* gene to be expressed throughout the wild-type eye disk^{11,12}, and when we induced *flb/flb* clones in the eye by mitotic recombination (*flb*^{2C82} and *Elp*^{BIRB3} were used), the mutant cells did not survive to the adult (data not shown). We do not know why *Elp* mutations predominantly affect the eye when amorphic mutations are lethal to the embryo. Possibly the increased *DER* activity in *Elp* is specific to the developing eye, or otherwise the eye might be especially sensitive to such an increase. It is also not known whether *Elp* mutations increase the expression of *DER* protein or enhance its activity, perhaps with respect to an eye-specific ligand or target. Another class of *DER* mutations, called *torpedo* alleles, preferentially affect the female germline, perturbing interactions between the follicle cells and the oocyte^{13,14}.

Enhanced activity of the *DER* could affect differentiation in a number of ways. In the present case, because even the earliest stages of differentiation were absent in *Elp/Elp* eye disks, the *DER* might be involved in the initial determination of ommatidia. In this model, differential *DER* activity controls where ommatidia will form, with high activity levels associated with the undifferentiated cells between preclusters, which later

give rise to pigment cells and bristles, as well as the remaining ommatidial cells; increased *DER* activity in *Elp/Elp* disks would then promote formation of these cells at the expense of pre-clusters. □

Received 26 April; accepted 31 May 1989.

- Tomlinson, A. *Development* **104**, 183–193 (1988).
- Ready, D. F. *Trends Neurosci.* **12**, 102–110 (1989).
- Tomlinson, A. T., Kimmel, B. E. & Rubin, G. M. *Cell* **55**, 771–784 (1988).
- Tomlinson, A. & Ready, D. F. *Devl Biol.* **120**, 366–376 (1987).
- Tomlinson, A., Bowtell, D. D. L., Hafen, E. & Rubin, G. M. *Cell* **51**, 143–150 (1987).
- O'Donnell, J., Boswell, R., Reynolds, T. & Mackay, W. *Genetics* **121**, 273–280 (1989).
- Müller, H. J. *Proc. VI Int. Cong. Genet.* **213**–255 (1932).
- Nüsslein-Volhard, C., Wieschaus, E. & Kluding, H. *Wilhelm Roux's Arch. dev. Biol.* **193**, 267–283 (1984).
- Price, J. V., Clifford, R. J. & Schupbach, T. *Cell* **56**, 1085–1092 (1989).
- Schejter, E. D. & Shilo, B.-Z. *Cell* **56**, 1093–1104 (1989).
- Schejter, E. D., Segal, D., Glazer, L. & Shilo, B.-Z. *Cell* **46**, 1091–1101 (1986).
- Kammermeyer, K. L. & Wadsworth, S. C. *Development* **100**, 201–210 (1987).
- Schupbach, T. *Cell* **49**, 699–707 (1987).
- Clifford, R. J. & Schupbach, T. *Genetics* (submitted).
- Lewis, E. B. & Bacher, F. *Drosoph. Inf. Serv.* **43**, 193 (1968).
- Lindsley, D. L. & Grell, E. H. *Genetic Variations of Drosophila melanogaster* (Carnegie Inst. Wash. Publ. 627, 1968).
- Campos-Ortega, J. A. & Gateff, E. A. *Wilhelm Roux's Arch. dev. Biol.* **179**, 373–393 (1976).
- Wieschaus, E. & Nüsslein-Volhard, C. in *Drosophila, A Practical Approach* (ed. Roberts, D. B.) 199–227 (IRL, Oxford/Washington, 1988).

ACKNOWLEDGEMENTS. We thank Drs J. O'Donnell, T. Schupbach and B.-Z. Shilo for unpublished information, Dr J. O'Donnell for mutant *Drosophila* stocks, Dr J. Brandenberger and the EM Laboratory, UC Berkeley for scanning electron microscopy, and Drs K. Anderson and A. Tomlinson for discussions. The manuscript was improved by criticism from Drs K. Anderson, E. Ferguson, C. Goodman, I. Hariharan and M. Simon. N.E.B. was supported by the Damon Runyon-Walter Winchell Cancer Research Fund.

Neurotransmitter inhibition of neuronal calcium currents by changes in channel voltage dependence

Bruce P. Bean

Department of Neurobiology, Harvard Medical School,
220 Longwood Avenue, Boston, Massachusetts 02115, USA

THE voltage-dependent calcium current of many neurons is depressed by transmitters such as noradrenaline^{1–3}, GABA^{4–6}, and kappa-opiate agonists^{7–12}. This modulation probably constitutes a major mechanism of presynaptic inhibition^{13,12}. Although recent work has implicated GTP-binding proteins in the mechanism of current inhibition^{13–19}, it is still unknown how the activation of those proteins alters the operation of the channels. In their initial description of the phenomenon¹, Dunlap and Fischbach proposed that noradrenaline acts by somehow reducing the number of functional calcium channels in the cell. By contrast with this hypothesis, I have found that inhibition of Ca²⁺ current is primarily due to a transmitter-induced change in the voltage-dependence with which channels are opened. Transmitters profoundly alter the voltage-dependence of channel activation, but there is little or no change in the number of functional channels activated by very large depolarizations. There is also little effect on the voltage-dependence of inactivation.

Figure 1 shows the effect of noradrenaline (NA) on Ca²⁺ currents in a bullfrog dorsal root ganglion (DRG) neuron. In this experiment, carried out on a freshly isolated cell body to ensure effective internal dialysis and optimal voltage control, NA inhibited the inward current elicited by small and moderate depolarizations by about 60% but had almost no effect on the outward current evoked by very large depolarizations. This outward current is carried by Cs⁺ flowing outward through the Ca²⁺ channel^{20,21}; it is missing when internal Cs⁺ is replaced by the tetraethyl ammonium ion (TEA), it is blocked by low

concentrations of La³⁺ (the records in Fig. 1 are La³⁺-subtracted); and it is reduced in parallel with inward currents by inactivation or application of ω -conotoxin.

Noradrenaline reduces the Ca²⁺-channel current without affecting its reversal potential (Fig. 1b). In the range –40 to +30 mV, currents are scaled down by NA with little apparent change in voltage dependence. The voltage-dependence of the action of NA becomes apparent only at more positive potentials and is most easily seen by plotting activation curves from the tail currents following various depolarizations (Fig. 1c). In the presence of NA, the activation curve becomes biphasic, with a component in the range –50 to +30 mV that is similar to the control activation curve, but about 60% smaller, and a new component of activation that saturates positive to +100 mV.

By contrast with its dramatic effects on channel activation, NA had only small effects on the voltage-dependence of inactivation (Fig. 2), shifting the midpoint of inactivation slightly in the depolarizing direction (the opposite direction to that which would produce inhibition).

In eight neurons in which 30 μ M NA reduced the peak Ca²⁺ current by more than 40%, as in Fig. 1, tail currents following steps to –20 mV were reduced by 54 $\pm$ 4% whereas those following steps to +130 to +150 mV were reduced by only 15 $\pm$ 3% (mean $\pm$ s.e.m). The effects of NA were blocked by phentolamine (30 μ M). Similarly dramatic voltage-dependent inhibition was seen with the κ -opiate agonist dynorphin A applied to rat DRG neurons (Fig. 3) and somatostatin applied to rat sympathetic neurons (data not shown).

These findings contradict previous reports that NA (ref. 1) and dynorphin^{7–8} do not alter the voltage-dependence of Ca²⁺-channel activation. There is no conflict, however, in the actual experimental data as these earlier recordings on cultured neurons resolved Ca²⁺-channel currents only up to $\sim$ +40 mV; in this limited voltage range, Ca²⁺ currents do, in fact, seem to be simply scaled down by the transmitter (Figs 1b, c and 3b).

These results show that the primary action of NA and dynorphin cannot be to reduce the number of functional Ca²⁺ channels. Almost all the channels are still functional following transmitter treatment, as shown by their ability to be opened within milliseconds by maximal depolarizations. Rather, a fraction of the channels seem to undergo a shift in voltage-dependence so

that they are no longer opened by small or moderate depolarizations but can still be opened by very large depolarizations. Another possibility, that the unitary conductance of the channels is reduced at moderate depolarizations, was ruled out by noise analysis experiments (see also ref. 4).

These results suggest the model shown in Fig. 4 in which it is assumed that channels can exist in two modes that are in equilibrium with one another. In the 'willing' mode, predominant in the absence of transmitter, channels can be opened effectively by moderate depolarizations. In the 'reluctant' mode, channels are not opened easily by small or moderate depolariz-

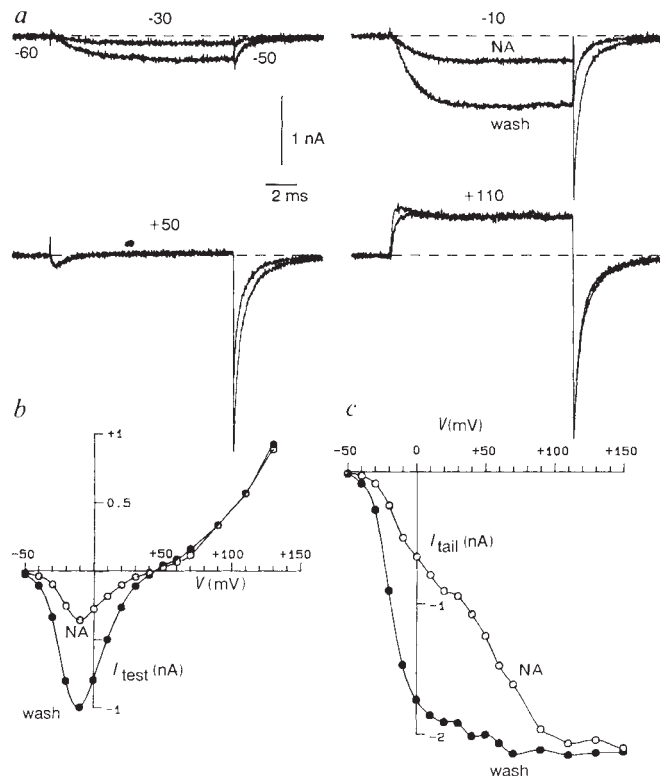


FIG. 1 Noradrenaline (NA) inhibition of Ca^{2+} -channel currents in a bullfrog neuron. *a*, Ca^{2+} -channel current at various potentials in a bullfrog DRG neuron in the presence and absence of $30 \mu\text{M}$ NA. After completing the series the cell was transferred to a solution containing $100 \mu\text{M}$ La^{3+} ; the currents remaining in La^{3+} (which were small) were subtracted to correct for non- Ca^{2+} -channel currents. Currents in NA were recorded 40–80 s after application and those labelled 'wash' were recorded 80–120 s after washing out the NA. The current at -10 mV declined from a control value of $1,165$ pA to 359 pA in NA and was restored to 998 pA after washing; currents in NA are compared to those after wash-out to restrict the comparison to the reversible effects of NA. *b*, Current during test pulse, averaged over the last 1 ms, as a function of voltage. *c*, Tail current, averaged over a 160 - μs interval beginning $160 \mu\text{s}$ after the end of the test pulse, as a function of the test voltage. Cell E72C; compensation for 1.2 of 1.6 -M Ω series resistance.

METHODS. Dorsal root ganglia from adult bullfrogs were incubated at 35°C for 45 min in Ca^{2+} -free Ringer's solution (in mM: 100 NaCl; 2.5 KCl; 5 MgCl_2 ; 10 glucose; 10 HEPES, pH adjusted to 7.4 with NaOH) containing 1 mg/ml $^{-1}$ collagenase (Sigma Type 1) and 5 mg/ml $^{-1}$ Dispase (Boehringer-Mannheim), then at 22°C for 45 min with Dispase alone. Single cells were released by trituration, stored at 5°C in Tyrode's solution (in mM: 150 NaCl; 4 KCl; 2 CaCl_2 ; 2 MgCl_2 ; 10 glucose; 10 HEPES, pH adjusted to 7.4 with NaOH), and used within 24 h. Cells were voltage-clamped at 22°C using patch pipettes³² with resistances of 0.3 – 2 M Ω . Internal solution (in mM): 126 Cs $^+$ glutamate; 9 EGTA; 4.5 MgCl_2 ; 9 HEPES; 3.6 MgATP; 14 creatine phosphate (Tris salt); 1 GTP (Na $^+$ salt); 50 U ml $^{-1}$ creatine phosphokinase; pH adjusted to 7.4 with CsOH. External solution (in mM): 3 BaCl_2 ; 154 TEA chloride; 10 HEPES; pH adjusted to 7.4 with TEA hydroxide.

ations but can be made to open by very large depolarizations. It is postulated that channels interconvert between the two modes and that transmitter action places a larger fraction in the reluctant mode. A quantitative formulation of the model can accurately account for the changes in the activation of Ca^{2+} channels produced by NA (Fig. 4*b, c*). In control, at any one time the vast majority (94%) of channels are in the willing mode (midpoint -15 mV). NA increases the fraction of channels in the reluctant mode (midpoint $+62$ mV) from 6% to 60% , resulting in a clearly biphasic activation curve.

Figure 4*d* shows an experiment which indicates that channels can interconvert between the willing and reluctant states rapidly. Using fast solution changes, NA was applied and removed during the time that Ca^{2+} channels were activated by a long depolarization. The inhibition of current by NA occurred within about 150 ms and the recovery from inhibition was similarly rapid, suggesting that conversion between the two modes can occur within this timescale.

This model can account for a commonly observed but mysterious phenomenon. With many transmitters, the Ca^{2+} current elicited in the first few msec is greatly reduced but then relaxes back towards the control current with a time course of the order 10 – 100 msec. Such changes in activation kinetics have been reported for noradrenaline^{4,22}, dopamine⁴, somatostatin^{23,24}, and enkephalin²⁵. As shown in Fig. 4*e* with somatostatin (records supplied by Dr. Dinah Sah), a current in the presence

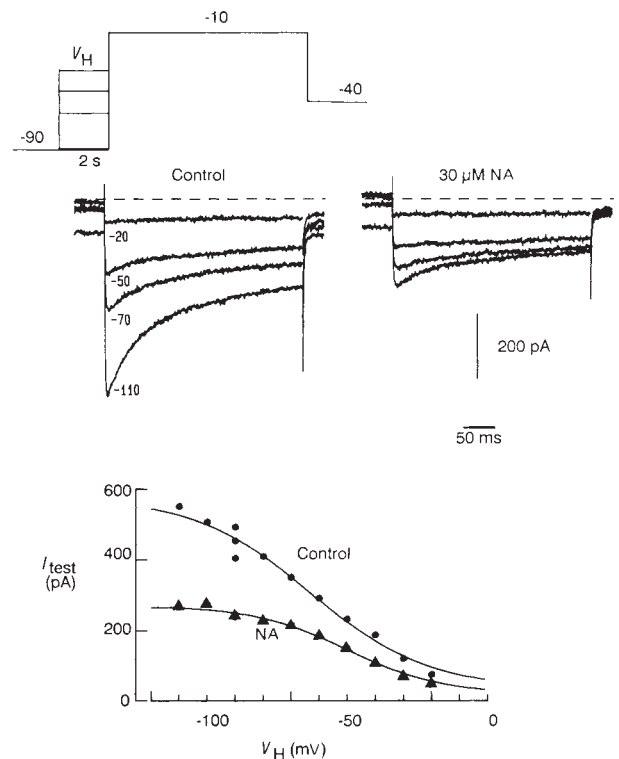


FIG. 2 Effect of NA on Ca^{2+} -channel inactivation in a bullfrog DRG neuron. A test pulse to -10 mV was given after different holding potentials were maintained for 2 s; the currents shown have been corrected for leak and capacitive currents obtained from a step from -90 to -98 mV. Top right: currents after $30 \mu\text{M}$ NA was applied for 0.5 – 2.5 min. Top left: currents recorded 1 – 3 min after NA was washed out. Bottom: peak test-pulse current versus holding potential. Solid curves are drawn according to $544/(1 + \exp((V_H - (-63))/21)) + 33$ for control points (circles) and $253/(1 + \exp((V_H - (-50))/17)) + 18$ for NA points (triangles). Conditions as in Fig. 1; cell E95C. Inactivation determined with this voltage protocol and these experimental conditions (internal EGTA, Ba^{2+} as charge carrier) is probably limited to voltage-dependent processes; possible effects of NA on current-dependent inactivation²² might not be evident.

of transmitter can relax nearly completely back to control levels. This behaviour is predicted by the model if equilibration between the willing and reluctant modes is sufficiently rapid. A moderate depolarization (to 0 mV, as in Fig. 4e) would produce in the first few msec almost complete activation of the channels in the willing mode but virtually no activation of channels in the reluctant mode. The depletion of channels in the closed, willing state would result in conversion of channels in the closed, reluctant state to the closed, willing state, followed by rapid activation. This would give rise to a slow phase of current activation, reflecting the conversion of reluctant to willing channels. For depolarizations large enough to open almost all willing channels, virtually all the initially reluctant channels would also eventually open.

The effects of NA and dynorphin are consistent with inhibition of a single predominant type of Ca^{2+} channel; the kinetics of both test pulse and tail currents remaining after NA treatment are nearly identical to control currents (Figs 1a and 3a). Furthermore, there is little or no difference in the kinetics of tail currents following moderate depolarizations (dramatically reduced by NA) and those following large depolarizations (little reduced by NA), suggesting that the voltage-dependent effects do not reflect selective actions on different channel types. The sensitivity of most of the Ca^{2+} current in bullfrog neurons to ω -conotoxin

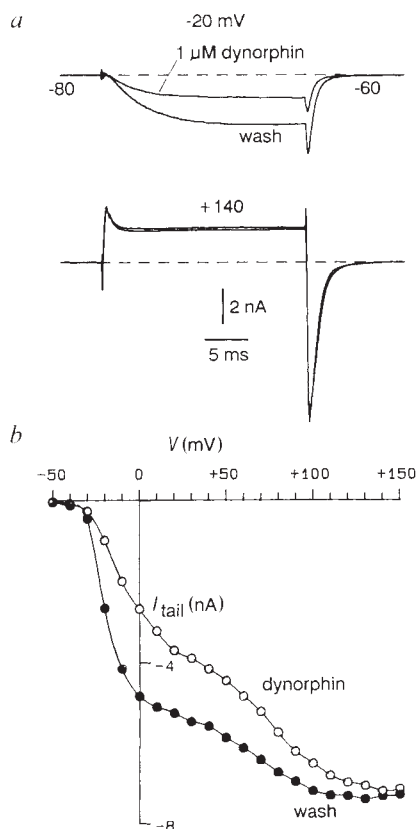


FIG. 3 Dynorphin A inhibition of Ca^{2+} -channel currents in a rat DRG neuron. *a*, As in Fig. 1, currents in dynorphin are compared with those after washout to focus on the reversible effect of dynorphin (there was a 20% decline in currents at all potentials between control and wash-out runs). *b*, Tail current, averaged over 400 μs beginning 680 μs after the end of the test pulse, versus test voltage. Cell W51D. Compensation for 2.9 of 4.2-M Ω m series resistance; 22 °C.

METHODS. The isolation technique was as for frog DRG neurons (Fig. 1 legend) except that the enzyme solution contained 1.5 mg/ml⁻¹ collagenase and 3 mg/ml⁻¹ Dispase in Ca^{2+} free Tyrode's solution. Internal solution (in mM): 115 CsCl; 5 MgATP; 5 creatine phosphate (Tris salt); 0.3 GTP (Tris salt); 10 BAPTA; and 10 HEPES; pH adjusted to 7.4 with CsOH. External solution (in mM): 3 BaCl₂; 160 TEA chloride; 10 HEPES; 3 μM tetrodotoxin; pH adjusted to 7.4 with TEA hydroxide.

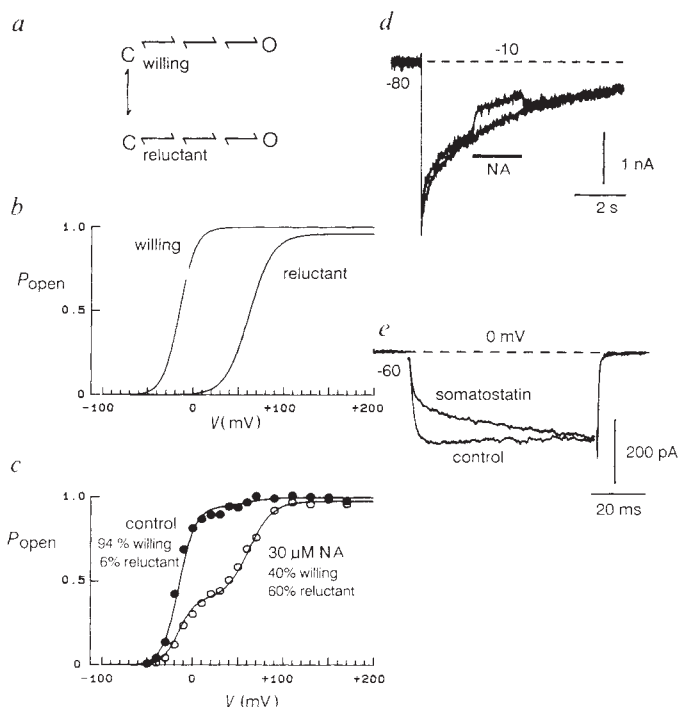


FIG. 4 A model for transmitter control of neuronal Ca^{2+} channels. *a*, It is assumed that channels exist in two modes, in equilibrium. *b*, P_{open} is the probability of a channel being open during a step to a given voltage. In one mode (willing), channels are opened readily by small depolarizations, whereas in the other (reluctant), channels require much larger depolarizations to open (and have a somewhat less steep voltage dependence and a slightly smaller maximal probability of being open). In each mode channels are assumed to have an activation curve given by a Boltzmann curve of the form $P_{\text{max}}/(1 + \exp(-(V - V_{0.5})/k))$, where P_{max} is the maximal probability of being open (1 for willing channels, 0.94 for reluctant channels); $V_{0.5}$, the midpoint, is -15 mV for willing channels and +62 mV for reluctant channels; and k is the slope factor (9 mV for willing channels, 13 mV for reluctant channels). *c*, With these parameters, the scheme accurately predicts the change in activation curve produced by 30 μM NA in a bullfrog sensory neuron (data from Fig. 1). *d*, Rapid kinetics of NA inhibition of Ca^{2+} -channel current in a bullfrog DRG neuron. Two successive currents elicited by long depolarizations to -10 mV are shown; during the second, the cell was transferred for 1.8 s to an external solution containing 30 μM NA. Methods and solutions as in Fig. 1 except the external solution contained 10 mM rather than 3 mM Ba²⁺. *e*, Slow activation phase of Ca^{2+} -channel current after application of somatostatin (1 μM) to a freshly dissociated spinal-cord neuron. Recording conditions as in Fig. 1 except that methanesulphonate replaced glutamate in the internal solution.

but not nitrendipine (unpublished results) indicates that N-type channels are predominant²⁶⁻²⁸. The current that is inhibited by NA inactivates in the range -100 to -20 mV, consistent with its classification as an N-type current^{26,27}.

Although stabilization of closed channels could result from phosphorylation by protein kinase C, which has been implicated in the action of NA^{17,18} (but see ref. 29), the kinetics of the action and reversal of NA seem surprisingly rapid for a phosphorylation-dephosphorylation mechanism. Another possibility is direct binding of G-protein subunits to the Ca^{2+} channel³⁰. The binding of a G-protein subunit to closed Ca^{2+} channels could interfere with the conformational change that opens the channel such that very large depolarizations are needed to drive the process.

Modulation of Ca^{2+} channels by a shift in voltage-dependence seems closely analogous to modulation of Ca^{2+} -activated K⁺ channels by a shift in Ca^{2+} -dependence³¹ and, more generally, to allosteric modulation of the substrate-dependence of enzymes. It will be interesting to see if altering voltage-dependence is a widespread mechanism by which transmitters regulate the activity of voltage-dependent channels. □

Received 2 February; accepted 2 June 1989.

- Dunlap, K. & Fischbach, G. D. *J. Physiol., Lond.* **317**, 519–535 (1981).
- Galvan, M. & Adams, P. R. *Brain Res.* **244**, 135–144 (1982).
- Forschner, P. & Oxford, G. S. *J. gen. Physiol.* **85**, 743–763 (1985).
- Marchetti, C., Carbone, E., & Lux, H. D. *Pflügers Arch.* **406**, 104–111 (1986).
- Dolphin, A. C. & Scott, R. H. *Brit. J. Pharm.* **88**, 213–220 (1986).
- Deisz, R. A. & Lux, H. D. *Neurosci. Lett.* **56**, 205–210 (1988).
- MacDonald, R. L. & Werz, M. A. *J. Physiol., Lond.* **377**, 237–249 (1986).
- Gross, R. A. & MacDonald, R. L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5469–5473 (1987).
- Miller, R. J. in *Structure and Function of the Slow Inward Calcium Channel* (eds Trigg, D. J. & Venter, J. C.), 161–246 (Liss, New York, 1987).
- Tsien, R. W., Lipscombe, D., Madison, D. V., Bley, K. R. & Fox, A. P. *Trends Neurosci.* **11**, 431–438 (1988).
- Bean, B. P. A. *Rev. Physiol.* **51**, 367–384 (1989).
- Tsien, R. W. in *Neuromodulation. The Biochemical Control of Neuronal Excitability* (eds Kaczmarek, L. K. & Levitan, I. B.), 206–242 (Oxford, New York, 1987).
- Holz, G. G., Rane, S. G. & Dunlap, K. *Nature* **319**, 670–672 (1986).
- Dolphin, A. C. & Scott, R. H. *J. Physiol., Lond.* **386**, 1–17 (1987).
- Hescheler, J., Rosenthal, W., Trautwein, W. & Shultz, G. *Nature* **325**, 445–447 (1987).
- Wanke, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 4313–4317 (1987).
- Rane, S. G. & Dunlap, K. 1986. *Proc. natn. Acad. Sci. U.S.A.* **83**, 184–188 (1986).
- Marchetti, C. & Brown, A. M. *Am. J. Physiol.* **23**, C206–C210 (1988).
- McFadzean, I. & Docherty, R. J. *Eur. J. Neurosci.* **1**, 141–147 (1989).
- Fenwick, E. M., Marty, A. & Neher, E. *J. Physiol., Lond.* **331**, 599–635 (1982).
- Lee, K. S. & Tsien, R. W. *Nature* **297**, 498–501 (1982).
- Docherty, R. J. & McFadzean, I. *Eur. J. Neurosci.* **1**, 132–140 (1989).
- Luini, A., Lewis, D., Guild, S., Schofield, G. G. & Weight, F. F. *J. Neurosci.* **6**, 3128–3132 (1986).
- Ikedo, S. R., Schofield, G. G. & Weight, F. F. *Neurosci. Lett.* **81**, 123–128 (1987).
- Tsunoo, A., Yoshi, M. & Narahashi, T. *Proc. natn. Acad. Sci.* **83**, 9832–9836 (1986).
- Fox, A. P., Nowycky, M. C. & Tsien, R. W. *J. Physiol., Lond.* **394**, 149–172 (1987).
- Fox, A. P., Nowycky, M. C. & Tsien, R. W. *J. Physiol., Lond.* **394**, 173–200 (1987).
- McCleskey, E. W. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 4327–4331 (1987).
- Hockberger, P., Toselli, M., Swandulla, D., & Lux, H. D. *Nature* **338**, 340–342 (1989).
- Yatani, A. *et al. Science* **238**, 1288–1292 (1987).
- Ewald, D., Williams, A. & Levitan, I. B. *Nature* **315**, 503–506 (1985).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch.* **391**, 85–100 (1981).

ACKNOWLEDGEMENTS. This work was supported by the NIH, the American Heart Association, and the Rita Allen Foundation. I thank Paul Ceelen for help in preparing cells and electrodes and Peter Hess, James Huettnner, Laura Regan, and Dinah Sah for discussions.

Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus

Detlev Khatchikian, Michaela Orlich & Rudolf Rott

Institut für Virologie, Justus-Liebig-Universität Giessen, 6300 Giessen, FRG

THE haemagglutinin glycoprotein HA of influenza viruses is responsible for the attachment of the virus to neuraminic acid-containing receptors at the cell surface and subsequent penetration by triggering fusion of the viral envelope with cellular membranes. To express full activity of the newly synthesized precursor, HA has to be modified by post-translational proteolytic cleavage into the polypeptides HA₁ and HA₂ by cellular enzymes. If proteases suitable for cleavage are not present in the host cell, the resulting virus particles are non-infectious^{1,2}. During adaptation of the apathogenic influenza virus A/turkey/Oregon/71 to chicken embryo cells, which are not permissive for HA cleavage³, we obtained an infectious virus variant with increased pathogenicity. Sequence analysis revealed that during adaptation 54 nucleotides were inserted into the HA gene; their sequence corresponds to a region of the 28S ribosomal RNA. This insertion is probably responsible for increased cleavability of HA, as well as for infectivity and pathogenicity of the adapted virus.

Adaptation of the influenza virus A/turkey/Oregon/71 (turkey virus) to chicken embryo cells (CEC) was obtained by a process of serial passages. Cells were infected with virus (less than one plaque-forming unit per cell) grown in embryonated eggs and were incubated in each passage for 48–72 h at 37 °C until suitable HA titres were found in the medium. The virus-containing medium used for the following passages was diluted 10–1,000 fold⁴. After 12 passages, a variant was found that grew efficiently under multiple-cycle conditions and produced

TABLE 1 Nucleotide and amino-acid changes in the HA gene of an influenza virus adaptation variant

Nucleotide position	Nucleotide change	Amino-acid changes
385	G → A	HA ₁ G → R (129)
455	G → A	HA ₁ S → N (152)
796	G → U	HA ₁ A → S (266)
1,035–1,036	insert of 54 nucleotides	(see Fig. 2)
1,060	G → C	HA ₂ G → R (354)

Extraction of viral RNA from purified virions and RNA sequence determination⁵ using synthetic primers of reverse transcription, were carried out as described¹². Numbers in brackets indicate the position of the amino acids in the mutant haemagglutinin.

plaques in CEC in the absence of trypsin. The capacity to grow in infectious form was not restricted to CEC; the adapted virus grew readily in canine (MDCK), hamster (BHK), bovine (MDBK) and monkey (Vero) cells, and was highly pathogenic for chickens. Whereas intramuscular infection with the apathogenic original turkey virus was asymptomatic, application of the variant to chickens led to a systemic, lethal infection with pathological signs of fowl plague.

Polyacrylamide gel electrophoresis revealed that the changes occurring in biological properties of the variant during adaptation were paralleled by cleavage of the HA in all cell types tested, as shown for CEC in Fig. 1. On the other hand, the HA of the unadapted virus was found exclusively in the uncleaved form. The haemagglutinin and the HA₁ polypeptide of the adapted variant exhibited a slight increase in relative molecular mass as compared with the original turkey virus.

To analyse the structural basis of the observed HA-processing, we sequenced the HA genes of the adapted and unadapted viruses by the chain termination method⁵. Four point mutations and one insertion of 54 nucleotides were detected in the HA gene of the variant (see Table 1). Previous data concerning the formation of the HA cleavage site² indicate that the point mutations in contrast to the insertion, should not influence HA cleavability. This is true even for the exchange of glycine by arginine at position 354, which may destabilize the structure of the mutant HA by inclusion of a charged amino acid into the hydrophobic HA₂ fusion peptide⁶. Sequence analyses of other variants, the passage histories of which were similar to those of the variant described here, however, revealed that this point mutation cannot influence the biological properties (unpublished data). Notably, the insertion was mapped in the region of the HA gene that encodes the C-terminus of HA₁, immediately adjacent to the cleavage site. This insertion corresponds in nucleotide sequence to a region in the 28S rRNA of mammalian cells^{7–9}. When compared with the corresponding sequence of human 28S rRNA (nucleotides 1,562–1,615)¹⁰, only one nucleotide was exchanged (Fig. 2). Unfortunately, no data are available on the nucleotide sequence of the 28S rRNA of chicken cells. In view of the strong conservation during evolution it can be assumed, however, that its structure is similar to that of the rRNA of mammalian cells.

Our findings provide substantial evidence that recombination between the rRNA of CEC and the HA gene furnishes the genetic basis of the insertion found in the adaptation variant. Although we cannot reconstruct the mechanism underlying this process, it is significant that immediately upstream from the insertion site eight nucleotides (AAAGACUA) of the plus-stranded HA-RNA (messenger RNA sense) are identical to the 3' border region of the inserted rRNA fragment. In the HA gene of the adapted virion, seven of these nucleotides are present in a palindrome, indicating that this sequence could have been a target site for the recombination (Fig. 2). As the insertion found in the HA gene (mRNA sense) of the variant showed the

opposite orientation to that of the rRNA (Fig. 2), we assume that the rRNA was used as a template during synthesis of positive-stranded viral RNA from the negative-sense template by polymerase jumping. Such a mechanism has been proposed to underlie recombinations between separate influenza virus genes. In this instance, the recombination in *trans* involved the viral RNA segments 1 and 3 only¹¹. Our findings, however, describe an insertion of a cellular RNA into the HA gene of an influenza virus.

It is probable that the inserted nucleotides are translated, leading to drastic alteration at the cleavage site of the HA molecule. This is indicated by the synthesis of a functional haemagglutinin and by the slightly higher relative molecular mass exhibited by the precursor HA and the polypeptide HA₁ of the adaptation variant compared with those of the original virus (Fig. 1). The addition of 18 amino acids at the connecting peptide could deform the cleavage site so as to render it susceptible to the action of ubiquitously occurring proteases. Consequently, the changes in the biological properties described for the CEC adapted variant could be due to this alteration at the HA cleavage site. The haemagglutinins of the naturally occurring pathogenic H7 avian influenza viruses sequenced so far are proteolytically activated in a broad range of cell types, and contain cleavage sites with short insertions of predominantly basic amino acids². It is possible that these inserts originated

from a similar recombination event between cellular and viral nucleic acids, as has been found in the CEC-adapted virus variant, and that such a mechanism could be of importance to the evolution of viruses in general. □

Received 3 April; accepted 2 June 1989.

1. Webster, R. G. & Rott, R. *Cell* **50**, 665–666 (1987).
2. Klenk, H.-D. & Rott, R. *Adv. Virus Res.* **34**, 247–281 (1988).
3. Klenk, H.-D., Rott, R., Orlich, M. & Blödmann, *Virology* **68**, 426–439 (1975).
4. Rott, R. *et al.* *EMBO J.* **3**, 3329–3332 (1984).
5. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
6. Gething, M.-J., Doms, R. W., York, D. & White, J. J. *Cell Biol.* **102**, 11–23 (1986).
7. Chan, Y. L., Olvera, J. & Wool, I. G. *Nucleic Acids Res.* **11**, 7819–7831 (1983).
8. Hassouna, U., Michot, B. & Bachellerie, J.-P. *Nucleic Acids Res.* **12**, 3563–3599 (1984).
9. Gonzales, I. L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 7660–7670 (1985).
10. Jerome, L. *et al.* *J. molec. Evol.* **24**, 236–251 (1987).
11. Fields, S. & Winter, G. *Cell* **28**, 303–313 (1982).
12. Daniels, R. S., Douglas, A. R., Skehel, J. J. & Wiley, D. C. *J. gen. Virol.* **64**, 1657–1662 (1983).
13. Bosch, F. X., Orlich, M., Klenk, H.-D. & Rott, R. *Virology* **95**, 197–207 (1979).

ACKNOWLEDGEMENTS. We thank Drs V. A. Erdmann and H. Niemann for discussions and acknowledge support from the Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie.

Cytolytic T-lymphocyte response to isolated class I H-2 proteins and influenza peptides

Kevin P. Kane*, Antonella Vitiello†, Linda A. Sherman† & Matthew F. Mescher*‡

* Division of Membrane Biology, Medical Biology Institute, La Jolla, California 92037, USA

† Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, California 92037, USA

T CELLS recognize antigenic peptides in the context of major histocompatibility complex (MHC) proteins. Peptide binding to class II MHC proteins^{1,2}, and T-cell recognition of these complexes at the functional level^{3–5} has been demonstrated. Although considerable evidence suggests that class I-restricted cytotoxic T lymphocytes (CTL) recognize class I-peptide complexes^{6–12}, this has not yet been directly demonstrated. Chen and Parham¹³ have recently detected a low level of direct binding of radiolabelled influenza peptides to class I HLA proteins, but the relevance of this binding to T-cell recognition remains uncertain. We report here that purified class I proteins pulsed with influenza peptides can trigger antigen-specific, TCR-mediated degranulation by CTL. Effective pulsing depends on both peptide concentration and time, and can occur within 60 minutes. These results provide strong support for the formation of an antigenic complex that is recognized by CTL in which peptide antigens are bound to isolated class I proteins.

We examined peptide recognition by H-2D^b-restricted CTL clones specific for the peptide corresponding to residues 365–380 of the A/PR/8/34 influenza nucleoprotein sequence NP(365–380). Whereas clones 4 and 36, are 'self'-restricted in that they are D^b-restricted and express D^b, two other NP-specific clones examined, 12C and 61C, were derived from radiation chimaeras and though D^b-restricted do not express D^b. All four clones lyse peptide-pulsed EL4 cells, which bear the appropriate D^b restriction element, but not peptide-pulsed RDM-4 (H-2^k) cells (not shown). Antigen-specific, TCR-dependent triggering of CTL can also be quantitated by measuring degranulation¹⁴, as assessed by the release of serine esterase into the medium¹⁵. Antigen-dependent release of serine esterase by the clones (Table 1) demonstrates the same specificity as that found for cytotoxicity. Furthermore, although self-restricted clones can respond to free peptide in the absence of targets, the clones obtained from

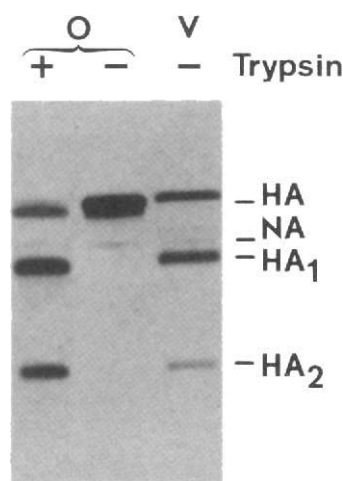


FIG. 1 Glycoproteins of influenza A/turkey/Oregon/71 virus (O) and its chicken-embryo cell adapted variant (V). Cells were labelled 4 h after infection with [³H]glucosamine. At ~20 h, cell-free virus was isolated, purified by adsorption to and elution from chicken erythrocytes and subsequent centrifugation. Original virus was treated (+) or non-treated (–) with 10 µg ml^{–1} trypsin for 15 min at 37 °C. SDS-polyacrylamide gel electrophoresis was carried out as previously described¹³.

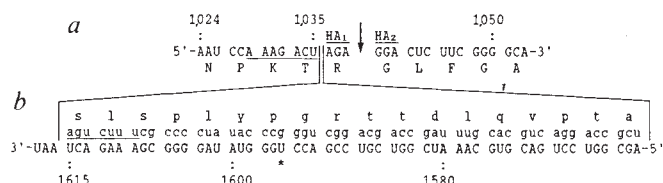


FIG. 2 Insertion of 54 nucleotides from the 28S ribosomal RNA into the HA gene of an avian influenza virus, adjacent to the HA cleavage site (arrow). The nucleotide sequence of the human rRNA homologous to the insertion (b) is taken from Jerome *et al.*¹⁰. (Asterisk, nucleotide exchange.) Sequence of the HA gene (a) is shown in the mRNA sense; sequence of the insert is marked by lower case letters. The palindrome at the 5' border region of the insert is underlined. The single-letter amino-acid code has been used for the protein sequences.

‡ To whom correspondence should be addressed

radiation chimaeras cannot (Table 1) as their lack of D^b, prevents them from presenting antigen to each other.

Purified class I proteins can be directly immobilized on plastic wells¹⁶ and alloantigen-specific CTL clones respond to the immobilized class I by specific, TCR-mediated release of serine esterase¹⁷. Ligand requirements for triggering clone 12C were examined using purified class I proteins and NP(365–380). Incubation of clone 12C with free peptide and immobilized D^b, but not K^k, resulted in degranulation (Fig. 1). Because these chimaera-derived CTL cannot present peptide to each other, this indicates unambiguously that responses result from the recognition of peptide and immobilized D^b. Clone 12C was also triggered when immobilized D^b, but not K^k, was incubated for 90 minutes with peptide and the wells then washed several times to remove free peptide (Fig. 1). These results strongly suggest that the NP peptide binds the immobilized D^b to form an antigenic complex recognized by clone 12C.

The specificity of the release of serine esterase in response to peptide-pulsed class I was further examined using a K^d-restricted CTL line (K^d/NP) specific for residues 147–158 of influenza NP. This line lyses target cells sensitized with NP(147–158R⁺), a modified form of NP(147–158) lacking an arginine and previously shown to sensitize target calls at low concentrations¹⁸. Although it responded to immobilized K^d, but not D^b or K^b, pulsed with NP(147–158R⁺) (Table 2), this line did not respond to any of the class I proteins pulsed with the NP(365–380) peptide. In the same experiment, clone 36 responded only to D^b pulsed with NP(365–380), whereas an allogeneic clone specific for K^b responded only to K^b, and this response was unaffected by peptide pulsing. Thus, responses to peptide-pulsed class I are clearly mediated by TCR recognition of antigen since they are completely specific for both the peptide and the restric-

TABLE 1 Specific degranulation of cloned NP(365–380)-specific, D^b-restricted CTL in response to free peptide and peptide-pulsed target cells

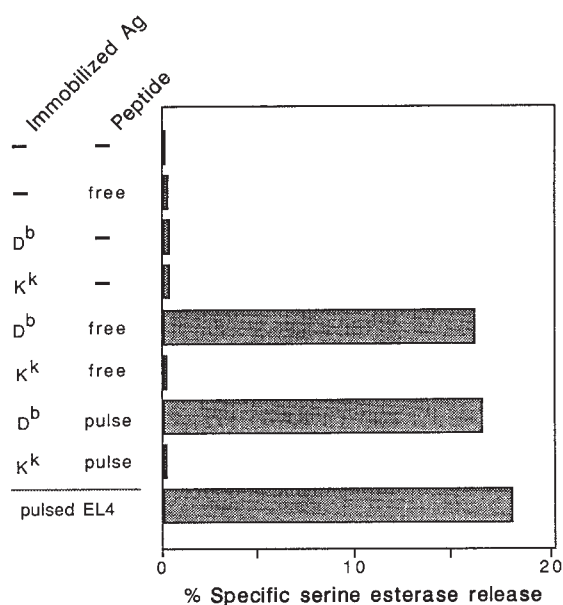
		Stimulus				
	Free		EL4	EL4	RDM4	RDM4
Clone	peptide	EL4	+free peptide	+pulsed peptide	+free peptide	+pulsed peptide
4	<u>0.483</u>	0.029	<u>0.411</u>	<u>0.506</u>	<u>0.467</u>	0.017
36	<u>0.318</u>	0.006	<u>0.257</u>	<u>0.242</u>	<u>0.327</u>	0.023
12C	0.012	0.021	<u>0.349</u>	<u>0.295</u>	0.031	0.032
61C	0.023	0.018	<u>0.237</u>	ND	ND	ND

Cloned CTL (10^5 per well) were incubated with the indicated stimulus for 4 h at 37°. Aliquots of supernatant were then removed and release of serine esterase activity determined. 'Free peptide' indicates that NP(365–380) was present at 28 μ M throughout the incubation. 'Pulsed peptide' indicates that EL4 (H2^b) or RDM4 (H2^b) were incubated with 28 μ M peptide in RPMI for 45 min, washed and then added in wells (2×10^5 per well) with CTL. See Fig. 1 legend for methods. Underlined values denote serine esterase activities several fold over spontaneous release values.

tion element. From our experiments examining responses of allogeneic CTL¹⁹, it seems probable that the interaction of CD8 with conserved class I determinants also contributes to the responses. It should be pointed out that the surface density of class I used here is similar to that on normal cells, and the CTL response levels are comparable with those obtained using peptide-pulsed targets as the stimulus (Fig. 1). Of eight peptide-specific CTL lines so far examined, two have failed to respond and six responded with appropriate specificity to immobilized class I determinants pulsed with peptide.

FIG. 1 Antigen-specific degranulation of clone 12C in response to immobilized H-2D^b pulsed with peptide. Class I antigens were immobilized at 0.2 μ g input per well and all wells then blocked for 30 min with 1 mg⁻¹ ml BSA. 'Free peptide' indicates that 28 μ M NP(365–380) was present throughout the assay. 'Pulse' indicates that 28 μ M peptide was added to wells following blocking with BSA and incubated for 1.5 h at 25 °C. Peptide was then removed and the wells were washed three times with FCS-containing medium. EL4 target cells were pulsed with 28 μ M peptide in RPMI with unbound peptide washed free with 2% FCS-RPMI. Clone 12C CTL were then added (10^5 per well) in 2% FCS-RPMI, and incubated for 4 h at 37 °C. Aliquots of supernatants were then removed and serine esterase activity determined. The specific esterase release is presented as the percentage of the total esterase activity in the CTL.

METHODS. Cloned CTL lines were derived from mice primed *in vivo* and maintained by weekly stimulation with influenza A/PR/8/34-infected spleen stimulator cells and recombinant interleukin-2. Clones were selected according to their ability to lyse virally infected or NP(365–380)-pulsed EL4 (H-2^b) target cells, and their specificity confirmed by their inability to lyse peptide-pulsed targets of other haplotypes (unpublished data and Vitiello *et al.*, manuscript submitted). Clones 4 and 36 were derived from C57BL/6 (H-2^b) responder mice. Clones 12C and 61C were derived from radiation chimaeras constructed by reconstitution of lethally irradiated (B10.D2 $\times$ C57BL/6)F1 mice with bone marrow derived from B10.D2 mice. A K^d-restricted, NP(147–158)-specific, long-term CTL line was established from cells of B10.D2 (H-2^d) responder mice, and its specificity confirmed in cytolytic assays. Peptides corresponding to influenza A/PR/8/34 nucleoprotein sequences (Multiple Peptide Systems, Inc.) were synthesized with a C-terminal amide. The NP(365–380) sequence was Ile-Ala-Ser-Asn-Glu-Asn-Met-Glu-Thr-Met-Glu-Ser-Ser-Thr-Leu-Glu and the NP(147–156R⁺) sequence was Thr-Tyr-Glu-Arg-Thr-Arg-Ala-Leu-Val-Thr-Glu, in which Arg 156 of the nucleoprotein sequence is deleted. Detergent-solubilized H-2D^b was purified by affinity chromatography on a 28-14-8 monoclonal antibody column, H-2K^k on an 11-4.1 mAb column, and H-2K^d on an M1/42 mAb column¹⁶. Antigens in 0.5% deoxycholate (DOC) were immobilized by diluting with PBS to 2 μ g⁻¹ ml and placing 0.1 ml into plastic microtitre plate wells (Falcon Micro Test III) to give 0.2 μ g per well. After 1.5 h, free H-2 antigen was removed by several washes with PBS and the wells then blocked with BSA. Degranulation assays were done using 10^5 effector CTLs in microtitre plate wells in a total volume of 0.1 ml of RPMI medium containing 2% FCS. At the end of the 37° incubation, 0.02 ml of supernatant was removed and assayed for



serine esterase activity by addition of a 0.18 ml N-benzoyloxycarbonyl-L-lysine thiobenzyl ester (BLT) reaction mixture, consisting of PBS, pH 7.4, with 2×10^{-4} M BLT and 2.2×10^{-4} M 5,5'-dithiobis (2-nitrobenzoic acid). The reaction was allowed to proceed for 20 min at 25° and the optical density (OD) at 412 nm then determined using an ELISA plate reader. Table 1 gives values of OD at 412 nm, after subtracting the value obtained for CTL incubated without stimulators (<0.05 OD units). The specific serine esterase release is calculated as $100 \times (E - S) / (T - S)$, where E is the experimental value, S is the esterase activity spontaneously released by CTL in the absence of stimulation, and T is the total activity extractable with TX-100. Spontaneous release was less than 0.05 OD units in all experiments. Data shown are the average of measurements of duplicate or triplicate wells, and replicate samples varied by less than 10% in all experiments.

Using clone 36, both the kinetics and the dependence of the class I interaction on peptide concentration were examined in more detail (Fig. 2). Peptide pulsing at a concentration of 56 μM for as short a time as 60–90 minutes resulted in a detectable response. A comparable response was evoked more slowly at lower peptide concentrations, and the maximal levels reached were lower. In all cases, pulsing for longer than 15 hours resulted in no further increase in response when CTL's were added. Assuming that response levels reflect the amount of peptide bound to class I, this suggests that equilibrium binding is reached within this time. The presence of other proteins does not prevent binding, as effective pulsing occurs either in phosphate-buffered

saline (Table 1), or in the presence of 2% serum (Table 2 and Fig. 2).

Chen and Parham¹³ have recently shown a low level (less than 0.3%) of binding of radiolabelled peptides to HLA proteins. This, and the results reported here strongly suggest that peptides can bind to native, mature class I proteins. More importantly, our results indicate that this can result in sufficient complex formation to allow triggering of CTL through the antigen-specific TCR. Thus, it seems probable that sensitization of targets by exogenous peptide could occur through the direct binding of peptide to the H-2 protein on the cell surface, as suggested by Hosken *et al.*²⁰. It may be the case that peptide binding to class I is limited by low affinity or the presence of bound peptide^{9,21}, and that very few complexes are required for triggering T cells. The approach described here allows a quantitative measurement of the recognition of peptide and isolated H-2 by CTL. The coupling of this approach with direct biochemical measurements of peptide binding to class I should enable the parameters of the peptide-class I interaction that contribute to the formation of a functionally active complex to be determined. □

TABLE 2 Specificity of anti-peptide and anti-alloantigen CTL lines for responses to immobilized class I proteins pulsed with peptides

Antigen	Effector CTL		
	C11 K ^b (allo)	C36 D ^b /NP(365–380)	K ^d /NP line K ^d /NP(147–158R ⁻)
D ^b	6	0	0
D ^b +NP(365–380)	5	25±1	0
D ^b +NP(147–158R ⁻)	4	0	0
K ^d	4	0	0
K ^d +NP(365–380)	3	0	0
K ^d +NP(147–158R ⁻)	2	0	18±1
K ^b	33±2	0	0
K ^b +NP(365–380)	37±1	0	0
K ^b +NP(147–158R ⁻)	36±2	0	0

Values represent % specific esterase release. H-2K^b, D^b and K^d were immobilized as described in Fig. 1 legend. Following immobilization, class I proteins were incubated in the absence or presence of the indicated peptides at 56 μM for 18 h at 37° in PBS, pH 7.4 with 2% FCS and 1% pen/strep. Free peptide was then removed by several washes with PBS, pH 7.4 plus 2% FCS. CTLs were then added at 10⁵ per well and incubated for 4 h at 37°; aliquots of supernatants were then removed and the serine esterase activity determined. Values are expressed as percentages of specific serine esterase release, as described in Fig. 1 legend. Standard deviations of triplicate values are shown for positive responses. Standard deviations on all other values were less than 2%. C11 is a cloned allogeneic CTL specific for H-2K^b. Clone 36 is specific for NP(365–380) and is D^b-restricted. K^d/NP is a long-term CTL line specific for NP(147–158R⁻) and is K^d-restricted.

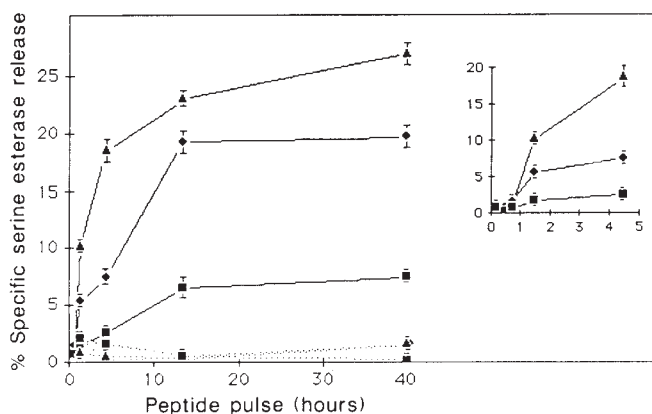


FIG. 2 Kinetics and dose dependence of NP(365–380) peptide pulsing of isolated D^b for CTL recognition. Isolated D^b (—) and K^b (·····) class I antigens were immobilized at 0.2 μg per well and unbound sites in wells were blocked with 1 mg ml^{-1} BSA. A pulse of peptide NP(365–380) at 56 μM (▲), 5.6 (●) and 0.56 (■) μM was done in 2% FCS–PBS (pH 7.4) containing 1% pen/strep at 37° for the indicated times. Peptide was removed and the wells were washed three times with FCS-containing medium. The NP(365–380)-D^b specific CTL clone 36 was then added (10⁵ per well) in 2% FCS–RPMI and incubated for 4 h at 37 °C. Spontaneous release by clone 36 was less than 5% of stimulated release. Error bars show the standard deviation obtained for measurements from triplicate wells. Inset shows the data from early time points replotted at a different scale.

Received April 10; accepted 2 June 1989.

- Babbitt, B., Allen, P., Matsueda, G., Haber, E. & Unanue, E. *Nature* **317**, 359–361 (1985).
- Buus, S. *et al.*, *Proc. natn. Acad. Sci. U.S.A.* **83**, 3968–3971 (1986).
- Buus, S., Sette, A., Colon, S., Jenis, D. & Grey, H. *Cell* **47**, 1071–1077 (1986).
- Watts, T. *J. Immunol.* **141**, 3708–3714 (1988).
- Sadegh-Nasseri, S. & McConnell, H. *Nature* **337**, 274–276 (1989).
- Townsend, A., Rothbard, J., Gotch, F., Bahadur, G., Wraith, D. & McMichael, A. *Cell* **44**, 959–968 (1986).
- Maryanski, J., Pala, P., Corradin, G., Jordan, B. & Cerottini, J.-C. *Nature* **324**, 578–579 (1986).
- Gotch, F., Rothbard, J., Howland, K., Townsend, A. & McMichael, A. *Nature* **326**, 881–882 (1987).
- Bjorkman, P. *et al.* *Nature* **329**, 506–512 (1987).
- Bjorkman, P. *et al.* *Nature* **329**, 513–518 (1987).
- Pala, P. *et al.* *J. Immunol.* **141**, 2289–2294 (1988).
- McMichael, A., Gotch, F., Santos-Aguado, J. & Strominger, J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9194–9198 (1988).
- Chen, B. & Parham, P. *Nature* **337**, 743–745 (1989).
- Pasternak, M., Verret, C., Liu, M. & Eisen, H. *Nature* **322**, 740–743 (1986).
- Takayama, H. *et al.* *J. Immunol.* **138**, 566–569 (1987).
- Kane, K., Champoux, P. & Mescher, M. *Molec. Immunol.* (in the press).
- Kane, K., Sherman, L. & Mescher, M. *J. Immunol.* **142**, 4153–4160 (1989).
- Bodmer, H., Pemberton, R., Rothbard, J. & Askanas, B. *Cell* **52**, 253–258 (1988).
- Goldstein, S. & Mescher, M. *J. Immunol.* **140**, 3707–3711 (1988).
- Hosken, N., Bevan, M. & Carbone, F. *J. Immunol.* **142**, 1079–1083 (1989).
- Parham, P. *Immunol. Today* **9**, 65–68 (1988).

ACKNOWLEDGEMENTS. This work was supported by the NIH (M.M. and L.A.S.), the American Cancer Society (M.M. and K.K.) and the NSF (L.A.S.).

Direct evidence that growth cones pull

Philip Lamoureux, Robert E. Buxbaum & Steven R. Heidemann

Department of Physiology, Michigan State University, E. Lansing, Michigan 48824-1101, USA

THERE is controversy over whether axonal elongation is the result of a pulling growth cone and the role of tension in axonal elongation^{1–4}. Earlier in this decade, the consensus was that axons or neurites elongated from tension generated by forward motility of the growth cone^{5,6}. It was presumed that contractile filopodia were the source of the tension moving the growth cone^{7,8}. But this view was challenged by experiments showing that neurites elongate, albeit abnormally, in the presence of cytochalasin, which inhibits growth-cone and filopodial movements⁹. Additionally, high resolution, video-enhanced observations of growth-cone activity argued against filopodial shortening as a source of tension, suggesting instead that an extrusion of cytoplasm rather than a pulling process, is the key event in neurite elongation^{1,3}. Studies of slow axonal transport¹⁰, however, indicate that much slower cytoskeletal

pushing underlies axonal elongation. We report here direct measurements of neurite force as a function of growth-cone advance which show that they are linearly related and accompanied by apparent neurite growth. No increase in force occurs in neurites whose growth cone fails to advance.

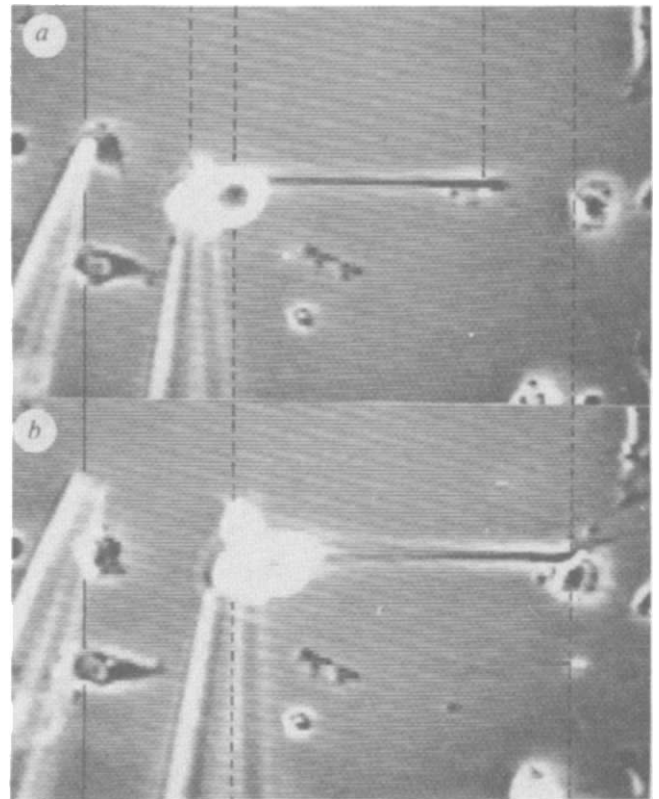
Glass needles of known compliance¹¹ were used to measure axial tension directly in neurites of cultured chick sensory neurons. As shown in Fig. 1, the calibrated glass needle was mounted in a micromanipulator next to a reference needle which moved only on movement of the micromanipulator. After attaching a calibrated needle to the cell body, the measurable distance between calibrated and reference needles provides a direct measure of axial force within the neurite. In these experiments the growth cone was the sole neuronal structure that remained attached to the culture dish and was allowed to advance for 'bouts' of 15 min to 3 h during which a time-lapse video-tape record was made. Measurements of directed growth-cone advance¹², and of the distance between calibrated and reference needles were taken at 3–5 min intervals from the videotape. Experiments on a given neuron were broken up into periods (bouts) that were separated from each other by more or less extensive micromanipulation of the neuron. In particular, after a period of active advance it was often necessary to prevent the growth cone from attaching to fibroblasts and other debris on the dish. This was done by pulling the growth cone backward to a clear area and waiting for its reattachment. A growth cone

manipulation of this type marked the end of a bout. The beginning of the next bout began when the growth cone was again attached to the dish, with a clear 'field' in front of it and the neurite axis aligned perpendicular to the needles. The sole micromanipulation adjustments during the course of a bout were small, vertical (see Fig. 1) adjustments to maintain the alignment of needles, neurite and growth cone. These were needed to avoid angular components of force along the neurite which would compromise the accuracy of the (horizontal) force measurements. To obtain estimates of axonal elongation, measurements of neurite length were made at the beginning and end of bouts at times of equivalent low neurite force, generally below 100 μ dynes, that is, when most of the neurite tension was eliminated by moving the neurite so as to slacken it. These differences in length cannot therefore be due to elastic stretching, as they are taken at equivalent tensions, and are approximate measures of elongation/growth. Ideally, we would have liked to measure neurite lengths at zero force before and after each bout; however we observed that neurites rapidly re-develop tension near zero force, as if to maintain some rest tension characteristic of cultured neurons¹¹.

The growth cones of six neurons advanced significantly in 17 bouts, the results of which are shown in Table 1 and Fig. 2. We found a strong temporal correlation between neurite tension increases and periods of growth cone advance: thus, as shown in Fig. 2e, f neurite tension remained constant when there was

FIG. 1 Neurite tension and growth-cone advance in chick sensory neurites. *a*, Video image near the beginning of bout 2A as reported in Table 1 and Fig. 2e–h. *b*, The same cell 75 min later in bout 2A. A solid line marks the position of the reference needle which fortuitously aligns with fibroblasts on the dish. The dashed lines show the change in calibrated needle deflection and growth-cone advance during the 75 min period. Magnification = 290 $\times$; 1 cm = 18 μ m.

METHODS. Embryonic chick sensory neurons were grown as previously described¹⁹. Glass needles were calibrated for their bending modulus as previously described¹¹, using a method similar to that of Yoneda²⁰. The bending modulus of the experimental needles varied between 5.6 and 10.6 μ dynes μ m⁻¹. The departure of these needles from hookean behaviour was estimated to be less than 4% based on their average working deflection and the fit of calibration data to first and second order equations. Needles were treated by immersion in 0.1% polylysine in phosphate buffered saline (PBS), followed by immersion in 1 mg ml⁻¹ collagen type IV (Sigma Chemical) in 0.1 M Tris pH 7.4, followed by immersion in 20 μ g ml⁻¹ laminin (Collaborative Research) in L-15 medium, all for 30 min at room temperature. This sequence of immersion was repeated with air drying between each step and the needle was stored. Immediately before use the needle was again treated with laminin as above. Needles were used repeatedly until they broke or no longer supported the cell body attachment. Recalibration of needles following use indicated no change in their modulus. The calibrated experimental needle was mounted in a micromanipulator along with a stiffer glass needle (60–120 μ dyn μ m⁻¹) that bore no load, serving as a reference for the movement of the micromanipulator, induced, for example, by thermal changes, and for the bending deflection of the calibrated needle. The micromanipulator was a modified Narishige hydraulic device model MO 203 with separate controls for the x, y and z axes. Neurons having neurites ~100 μ m long and with a clear field before the growth cone were oriented so that the neurite was horizontal on the television monitor. These cells were then manipulated to get the cell body to attach to calibrated needle without detaching the growth cone with the neurite oriented perpendicular to the axis of the calibrated needle. To maintain the cells at physiological temperature while minimizing thermal changes in micromanipulator position, the entire laboratory room was maintained at 37 $\pm$ 1 $^{\circ}$. Local environmental regulators were not used as they caused thermal movements of the micromanipulator indicated by instability in the reference-needle position. Results were used only from bouts where the reference needle spontaneously moved less than 3 μ m: in most cases the reference needle did not move measurably. After attachment to the cell body, the needles were raised slightly so that only the growth cone remained attached to the substratum. The growth cone was allowed to advance (or not) for periods of 15 min to 3 h during which a time lapse video-tape record was made. During this time, only small 'vertical' movements of the micromanipulator were imposed on the system to maintain the neurite and growth



cone perpendicular to the axis of the calibrated needle. An example is the vertical shift in the position of needles between panels *a* and *b*. Measurements of growth-cone advance by the method of Katz *et al.*¹² used the neurite 'neck,' the very proximal end of the growth cone, as the point of reference. This measurement, needle-separation and reference-needle position were recorded every 3–5 min from a replay of the video tape on a video screen with a 1-mm gridwork overlay. The gridwork was calibrated using an image of a stage micrometer. We estimate the reproducibility error of video measurements to be about 10% based on a comparison of the data reported here and measurements taken by a naive student.

FIG. 2 Growth-cone advance, neurite tension and elongation in two bouts from two neurons. *a-d*, Present data from bout 1B and *e-h* from bout 2A, reported in Table 1. *a* and *e*, Growth-cone advance as a function of time (min). *b* and *f*, Neurite tension as a function of time (min). *c* and *g*, Net elongation (see text) of the neurite shown as a straight line representing the length change during the time of the bout. *d* and *h*, Neurite tension as a function of growth-cone advance, with correlation coefficients, R , shown. See text for explanation of the non-zero force intercept.

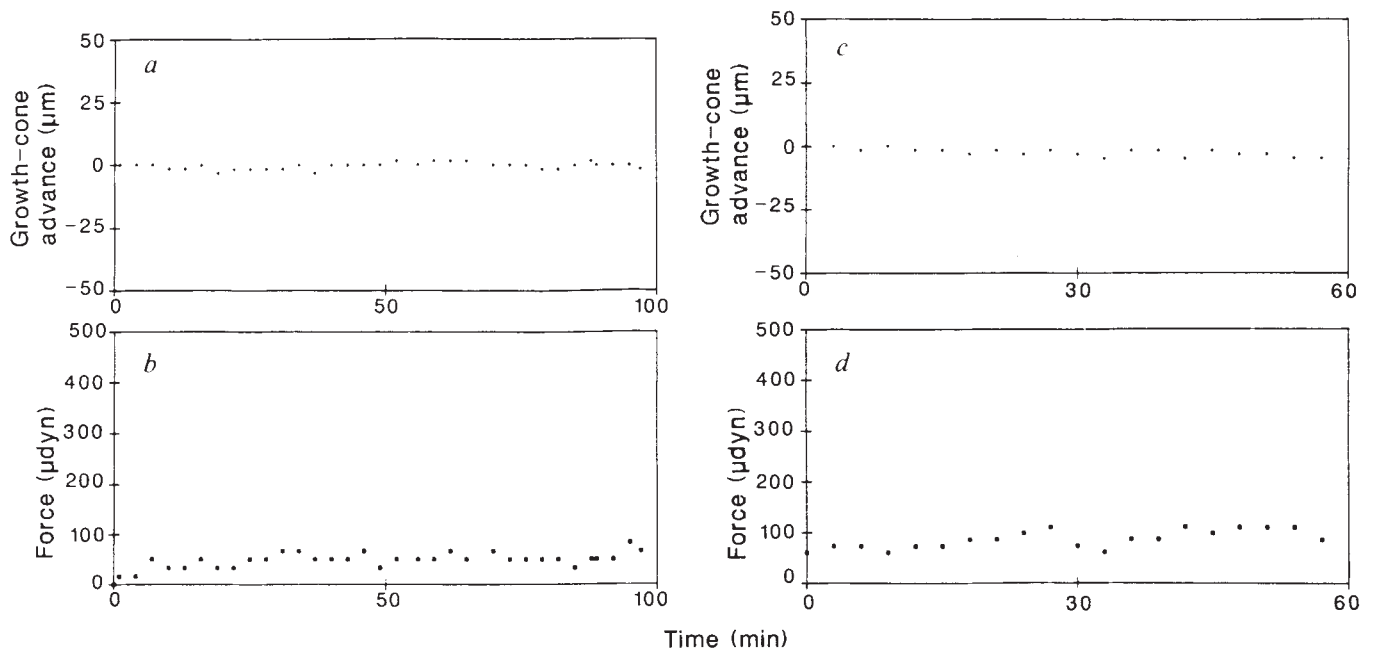
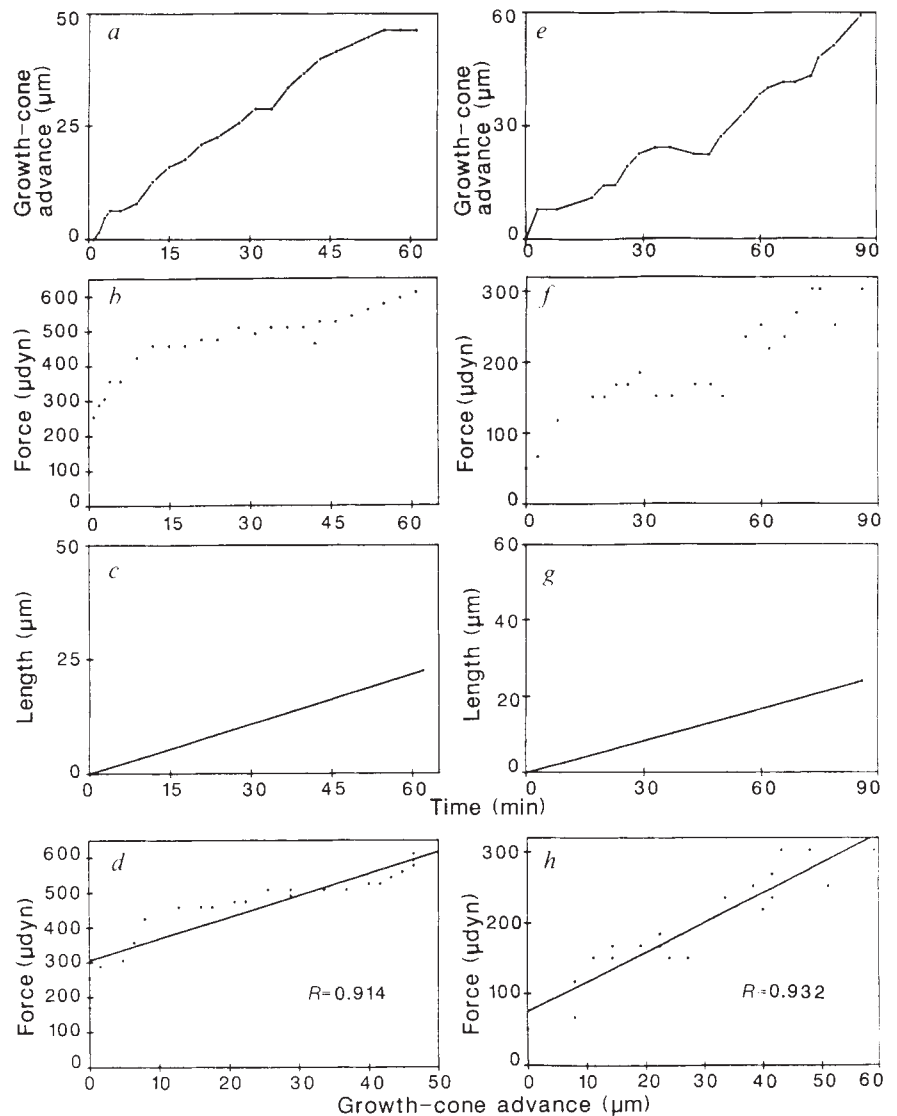


FIG. 3 Lack of neurite tension increase in the absence of growth-cone advance. *a* and *b* present data from one neurite, *c* and *d* data from another.

a and *c*, Growth-cone advance as a function of time. *b* and *d*, Neurite tension as a function of time.

TABLE 1 Summary of growth-cone advance observations

Neurite	Bout	Duration (min)	Growth-cone advance (μm)	Tension increase (μdyn)	Correlation (r)	Net elongation (growth) (μm)
1	A	192	83	509	0.927	48
	B	62	46	431	0.914	22
	C	58	27	305	0.871	10
	D	34	24	220	0.956	5
2	A	88	59	269	0.932	24
	B	16	28	218	0.966	8
	C	22	42	152	0.972	11
	D	53	64	185	0.924	16
3	A	17	14	146	0.994	3
	B	48	29	130	0.895	—
4	A	18	24	65	0.959	16
	B	19	24	65	0.912	10
	C	77	59	130	0.954	46
5	A	41	16	46	0.934	—
6	A	231	56	149	0.889	38
	B	22	10	82	0.969	3
	C	70	27	217	0.943	-3

relatively little growth-cone advance between 30 and 50 min, but increased sharply between 50 and 70 min as the growth cone advanced rapidly. Neurite tension was strongly correlated with growth-cone advance and correlation coefficients were 0.87 or greater (Table 1). The linear relationship between growth-cone advance (strain) and tension (stress) suggests that a substantial portion of the advance was accommodated by elastic stretching of the neurite. Growth cone advance, however, was also accompanied by net neurite elongation (growth): the length of neurites corrected for elastic stretching increased in 14 of 15 bouts in which we were able to make measurements.

The relative contributions of elastic stretching and elongation varied from bout to bout. Elongation (growth) accounted for 21–78% of the advance, an average of 47% considering all 14 bouts together. Although this large fraction of growth-cone advance did not contribute to the increase in elastic force, growth-cone advance nevertheless correlated with the tension increase (coefficients 0.9). This indicates that growth/elongation is also highly correlated with tension, otherwise the overall correlation would be reduced. For example, a neurite extension that was 50% stretching and 50% elongation, where tension correlates to overall advance at 0.9, must have at least a 0.8 correlation between tension and elongation.

Obviously, the increase in force with growth-cone advance cannot continue indefinitely and we were surprised by the length of the period (up to 90 min) during which force correlated linearly with growth-cone advance. In some experiments (for example 1B, 6A) the force did level off in spite of growth-cone advance. Although Bray's¹³ description of towed growth suggests a more rapid levelling of tension as growth fully accommodates the advance, our data and preliminary measurements of tension in towed neurites show linear force increases for a significant period during neurite terminal advance. Because the relationship between tension and mass addition, which would presumably dissipate the tension, remains poorly understood further work will be necessary to interpret this aspect.

That growth-cone advance does not occur continuously¹² was shown in eight experimental neurons (eight bouts) typical examples of which are given in Fig. 3. Significantly, the neurite tension remained nearly constant in all eight cases and the neurites appeared normal. Indeed, one neuron provided both a bout with, and a bout without, growth cone advance. The lack of increase in tension in the absence of growth-cone advance indicates that the neurite shaft does not contribute to tension increases during growth cone advance. As already noted,

however, neurites rapidly redevelop force when completely slackened. This tension increase is presumably generated within the neurite shaft, and is observed only during the first few minutes (following micromanipulations before each bout) as shown by the non-zero intercepts in Figs 2 and 3. Cell bodies did not move on the needle in any of these experiments and thus did not contribute directly to tension increases.

These data provide direct evidence that growth cones pull their neurite forward. Growth-cone advance and increases in tension were strongly correlated and both were correlated in time. Additionally elongation/growth seems to be correlated with tension. We conclude that a pulling growth cone provides an important stimulus for growth under standard culture conditions. A pulling growth cone, however, is not an absolute requirement for growth^{4,9,14}; we have previously suggested how the mechanical relationships of the axonal cytoskeletal elements might produce growth, under different culture conditions, with net compression on the neurite^{11,15}. The mechanism of tension generation by the growth cone remains problematic. Our time-lapse observations of chick sensory growth cones (unpublished) are similar to those of recently published reports^{1,3} in showing that filopodial contraction does not underlie growth-cone advance. The retrograde movement of actin in the growth cone has recently been proposed as a mechanism for tension generation^{16–18}. □

Received 2 March; accepted 6 June 1989.

1. Aletta, J. M. & Greene, L. A. *J. Neurosci.* **8**, 1425–1435 (1988).
2. Bray, D. *Trends Neurosci.* **10**, 431–434 (1987).
3. Goldberg, D. J. & Burmeister, D. W. *J. Cell Biol.* **103**, 1921–1931 (1986).
4. Letourneau, P. C., Shattuck, T. A. & Ressler, A. H. *Cell Motil. Cytoskeleton* **8**, 193–209 (1987).
5. Landis, S. C. A. *Rev. Physiol.* **45**, 567–580 (1983).
6. Letourneau, P. C. in: *Neuronal Development* (ed. Spitzer, N. C.) (Plenum, New York, 1982).
7. Bray, D. in: *Cell Behavior* (ed. Bellair, R., Curtis, A. & Dunn, G.) (Cambridge University Press, 1982).
8. Trinkaus, J. P. *J. Neurosci. Res.* **13**, 1–19 (1985).
9. Marsh, L. & Letourneau, P. C. *J. Cell Biol.* **99**, 2061–2067 (1984).
10. Lasek, R. J. *J. Cell Sci. suppl.* **5**, 161–179 (1986).
11. Dennerli, T. J., Joshi H. C., Steel, V. L., Buxbaum, R. E. & Heidemann, S. R. *J. Cell Biol.* **107**, 665–674 (1988).
12. Katz, M. J., George, E. B. & Gilbert, L. J. *Cell Motil.* **4**, 351–370 (1984).
13. Bray, D. *Dev. Biol.* **102**, 379–389 (1984).
14. Spero, D. A. & Roisen, F. J. *Dev. Brain Res.* **23**, 155–159 (1985).
15. Buxbaum, R. E. & Heidemann, S. R. *J. theor. Biol.* **134**, 379–390 (1988).
16. Bray, D. & White, J. G. *Science* **239**, 883–888 (1988).
17. Forscher, P. & Smith, S. J. *J. Cell Biol.* **107**, 1505–1516 (1988).
18. Mitchison, T. & Kirschner, M. *Neuron* **1**, 761–772 (1988).
19. Baas, P. W. & Heidemann, S. R. *J. Cell Biol.* **103**, 917–927 (1986).
20. Yoneda, M. *J. exp. Biol.* **37**, 461–468 (1960).

ACKNOWLEDGEMENTS. This work was supported by the NSF and NIH. We thank Tim Dennerli for discussions.

COUP transcription factor is a member of the steroid receptor superfamily

Lee-Ho Wang, Sophia Y. Tsai, Richard G. Cook*, Wanda G. Beattie, Ming-Jer Tsai & Bert W. O'Malley

Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030, USA

* Department of Microbiology and Immunology, Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030, USA

THE COUP (chicken ovalbumin upstream promoter) transcription factor (COUP-TF) exists in a number of different tissues and is essential for expression of the chicken ovalbumin gene¹⁻³. It binds to the ovalbumin promoter and, in conjunction with a second protein (S300-II), stimulates initiation of transcription *in vitro*. COUP-TF also binds specifically to the rat insulin promoter element^{4,5}, although the two binding sites share little sequence similarity. Here we report the isolation of a human complementary DNA clone encoding COUP-TF. Comparison of the amino-acid sequence of COUP-TF with known sequences reveals that it is a member of the steroid/thyroid hormone/vitamin receptor superfamily⁶. Consequently, it is the first member of this family that has been shown to function in a cell-free transcription system^{7,8}. We conclude that this superfamily of gene regulators contains proteins which bind and activate distal promoter elements of eukaryotic genes.

To obtain a high yield of purified COUP-TF HeLa nuclear extracts were applied three times to a sequence-specific DNA affinity column. After the third consecutive step a protein sample revealed five major polypeptide bands of molecular mass (M_r) 53, 48, 46, 44 and 43K (Fig. 1b). Polyclonal antibodies were raised in rabbits against the native proteins prepared from the third passage of the affinity column. On interaction with polyclonal antibody (Fig. 1a) the migration of C1 and C2 (COUP-TF-DNA complexes) in a native polyacrylamide gel indicated the formation of an antibody *en*-COUP-TF complex. This mobility change was not found if COUP-TF was incubated with either pre-immune serum or control antiserum to heat-shock protein. Western blots of HeLa-cell extracts showed that the antiserum interacted predominantly with the 46 and 48K polypeptide doublet (Fig. 1b). This antibody was therefore considered suitable for screening a HeLa cell cDNA expression library.

To clone the gene encoding the COUP-TF, a library of randomly primed and oligo(dT) primed cDNAs was constructed in λ gt11 using poly(A)⁺ messenger RNA from HeLa cells⁹. One clone cross-reacted with both antibody and concatenated binding site¹⁰, indicating that it contained an appropriate DNA-binding domain and an epitope(s) for the antibody. We isolated this clone for testing for specific binding to the authentic COUP sequence.

Lysogenic cell extracts from this recombinant plaque were incubated with the DNA fragment containing the COUP sequence and assayed in a gel retardation assay. Several protein-DNA complexes were formed when IPTG-induced (isopropyl β -D-thiogalactopyranoside) cell extracts were used but were not seen in uninduced extracts (Fig. 2a). Binding specificity was also demonstrated by competition analysis. The formation of complexes could be abolished using an unlabelled oligonucleotide corresponding to the wild-type COUP sequence, whereas a mutant oligonucleotide failed to compete at 50-fold or 100-fold molar excess (Fig. 2a). Furthermore, methylation interference analysis revealed that the purine contact sites of the fusion protein with the COUP sequence were identical to those found using HeLa COUP-TF (Fig. 2b). These results support the conclusion that the bacterial fusion protein binds exactly as predicted to the COUP-DNA sequence.

The gene insert of this clone, 1,513 nucleotides long, contained an open reading frame of 1,254 bases. The nucleotide and deduced amino-acid sequences are shown in Fig. 3. Comparison of amino-acid sequences of five different polypeptide fragments from the 46 and 48K polypeptides obtained after CNBr cleavage, together with the amino-acid sequence deduced from the cDNA clone, confirmed the correct reading frame and were evidence of the authenticity of the cDNA clone.

To investigate any relationship with known proteins, we used the PSEARCH program¹¹ in a computer-assisted search for sequences related to COUP-TF in the EMBL/GenBank libraries. Surprisingly, the amino-acid sequence deduced from the COUP-TF cDNA revealed significant similarities to members of the steroid-hormone receptor superfamily of genes (Fig. 4). We

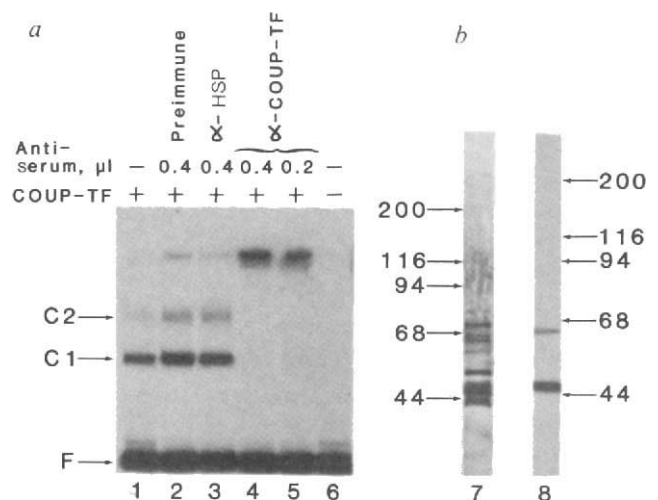


FIG. 1 Characterization of antiserum to COUP TF by gel-retardation and western-blot analyses. *a*, Partially purified HeLa COUP-TF was assayed in a gel-retardation assay with pre-immune serum (lane 2), antiserum to heat-shock protein (HSP) (lane 3) and antiserum to COUP-TF (lanes 4 and 5); a DNA fragment spanning -269 to -44 of the ovalbumin gene (containing COUP sequence) was used as the probe¹. Lane 6 contained probe alone. Binding reactions, as previously described¹, were carried out in the presence of 2 μ g *Hinf*I-digested pBR322 as a non-specific competitor. We demonstrated previously that four of the polypeptides (43-48K) from an SDS-PAGE gel, after denaturation and renaturation, could bind the COUP sequence specifically. In addition, some minor polypeptides clustering about the molecular mass of 68K were observed. These 68K polypeptides also bound specifically to the COUP sequence, and the patterns of the indirect DNase I footprinting of both 43-48K and 68K polypeptides were indistinguishable (data not shown). The bands corresponding to the free probe, the 43-48K polypeptide-DNA complex and the 68K polypeptide-DNA complex are indicated as F, C1 and C2, respectively. *b* In lane 7, a 0.42 μ g aliquot of the COUP-TF from the third passage of affinity column was precipitated with trichloroacetic acid, subjected to an SDS-10% PAGE gel and stained with silver. Values of $M_r \times 10^{-3}$ protein markers are indicated. In lane 8, a 140 μ g aliquot of HeLa whole cell extract was run separately on an SDS-10% PAGE gel and transferred to an Immobilon PVDF Membrane²⁰. After blocking with 3% non-fat milk, the membrane was incubated with a 500-fold dilution of COUP-TF antiserum. ¹²⁵I-labelled protein A was used to visualize the signal. **METHODS.** Nuclear extracts of HeLa cells (Invitron, Berkeley, MO) were prepared as described previously²¹. Typically, 120 ml nuclear extract were loaded at 0.25 M NaCl on a 5-ml COUP-specific DNA affinity column². NaCl eluate (0.6M) containing COUP-TF was adjusted to 0.3 M in NaCl and incubated with 300 μ g ml⁻¹ *Hinf*I-digested pBR322 for 30 min before application to 2 ml affinity resin. Repeat applications to the affinity column differed only on the third cycle when 150 μ g ml⁻¹ non-specific DNA competitor and 1 ml affinity column resin were used, to yield ~3.5 μ g COUP-TF. Forty μ g proteins were concentrated by heparin-Sepharose chromatography and lyophilization. Two doses of 10 μ g purified COUP-TF were injected into two New Zealand White rabbits. The first was given intradermally in complete Freund's adjuvant and a booster was injected subcutaneously with incomplete adjuvant four-weeks later. Antiserum was tested without purification.

identified a 66-amino acid Zn-finger motif in which all 20 invariant amino acids were conserved and 11 out of 12 conserved residues were identical¹². The only amino acid that differs from the conserved Zn-finger sequence is a glutamine, which replaces a conserved lysine in the second finger. A schematic amino-acid comparison of COUP-TF with the human steroid-hormone receptor superfamily is shown in Fig. 4a. The Zn-finger motif is shown as region I. A consensus sequence was derived for this region using members of this family (Fig. 4b).

In addition to the Zn-finger motif, there are two new distinct regions (shown as II and III in Fig. 4a), which reveal significant similarity among the members of this superfamily. Both regions

are in the ligand-binding domain; sequence comparisons, together with a consensus sequence, are shown in Fig. 4c and d. We observed that COUP-TF contains homologies to hERR-1 and hERR-2 (ref. 13) over 100 amino-acid residues, including two regions designated II and III. The significance of these two newly identified conserved regions (II and III) in the steroid-receptor family is unknown. But the strong conservation of these sequences indicates that these regions could have a function in this group of transcription factors¹⁴⁻¹⁷.

A feature of steroid receptors is their capacity to bind hormonal ligands. To date, we have not identified a ligand which associates specifically with the COUP-TF. It is possible that

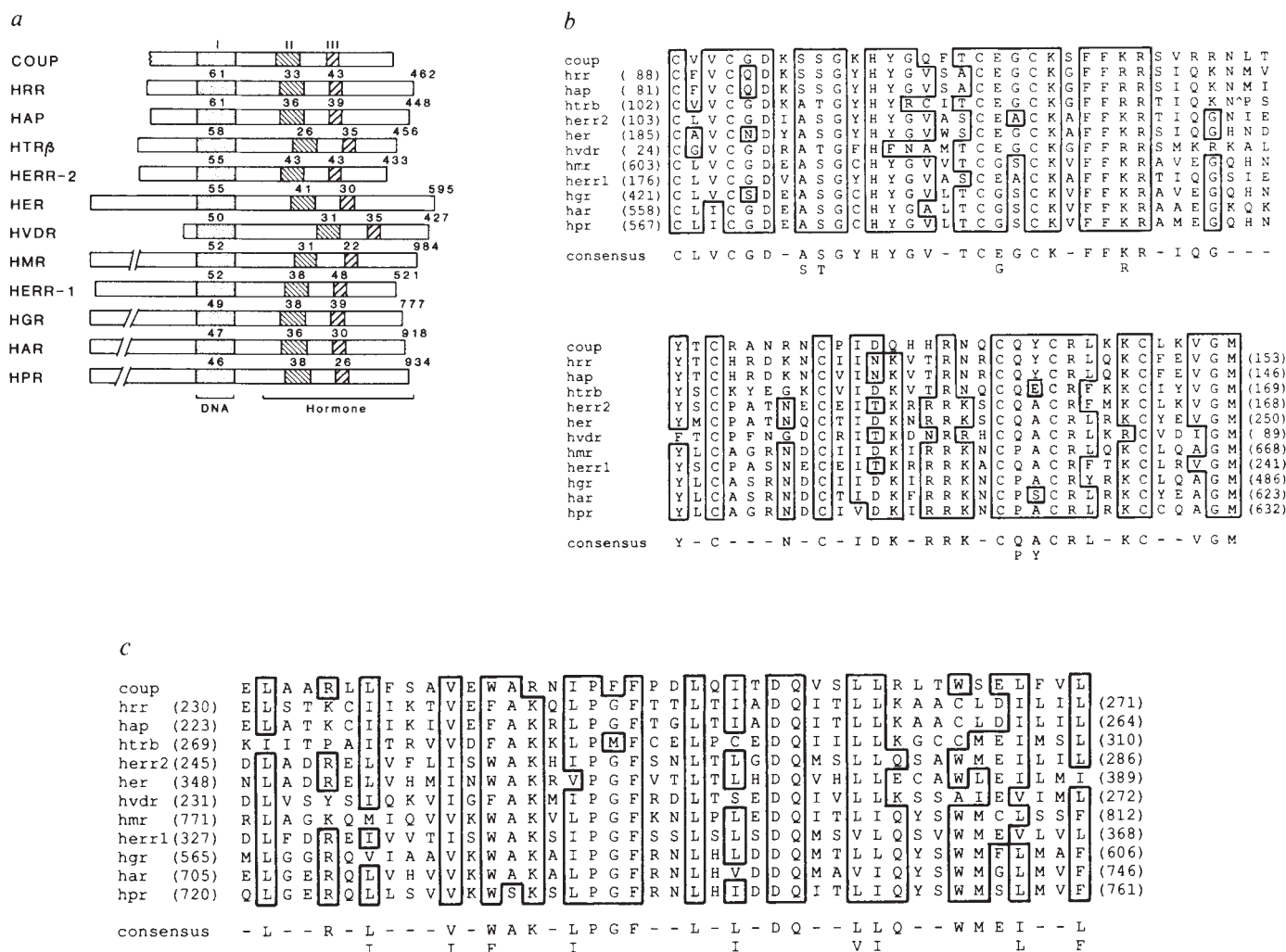


FIG. 4 Schematic amino-acid comparison of members of the human steroid-hormone receptor superfamily. *a*, Alignment of amino-acid sequences on the basis of maximum similarity identified three regions shown as I, II and III. The number of amino-acid residues in these three regions are: I, 66; II, 42; and III, 23. The numbers above each box indicate the percentage of similarity in each region as compared to the corresponding region of COUP-TF. The number at the C-terminal of individual receptors represents the total amino-acid residues. The DNA and hormone-binding domains are indicated. Specific references are: RR (refs 24, 25), retinoic-acid receptor; HAP (ref. 26) TR β (ref. 27), thyroid receptor; ERR1 and ERR2 (ref. 13), oestrogen receptor-related (cDNAs) 1 or 2; ER (ref. 28), oestrogen receptor; VDR (ref. 29), vitamin-D₃ receptor; MR (ref. 30), mineralocorticoid receptor; GR (ref. 31), glucocorticoid receptor; AR (refs 32, 33), androgen receptor; and PR (ref. 34), progesterone receptor. The prefix H denotes human in all cases. Amino-acid sequence alignment of the Zn-finger motif (I), region II and region III of the COUP-TF and members of human steroid-hormone receptor superfamily are shown in *b*, *c* and *d*, respectively. The numbers in brackets represent the

positions of amino-acid residues in the individual receptors. Boxed amino acids are those identified to be the consensus amino-acid sequences. In htrb there are two additional amino-acid residues (Leu and His) at position indicated by Δ in panel *b*.

COUP-TF can function in the absence of any ligand. As COUP-TF can be activated transcriptionally by S300-II (ref. 8), it is conceivable that S300-II acts in place of a ligand. Because of the conservation of amino acids in the C-terminal region within the steroid-receptor superfamily, however, we hypothesize that this promoter regulatory protein is activated by a non-steroidal ligand which is yet to be discovered.

Upstream from the Zn-finger motif, the N-terminal portion of COUP-TF is fairly rich in G·C base pairs (>85%). The 78 amino-acid residues in this region include 12 proline and 12 glutamine residues. A Gln-rich region has been reported to be the transactivation domain of the Spl transcription factor¹⁸. This N-terminal region of COUP-TF may also be involved in transcriptional activation.

Here, we present a sequence derived from a clone of the COUP-TF and have identified it as a new member of the steroid-receptor superfamily of genes. COUP-TF represents the first instance in which a protein that functions as a promoter transcription factor has been classified as a member of the steroid/thyroid-hormone/vitamin receptor superfamily. The COUP-TF sequence was published recently in a clone designated ear-3, obtained by cross hybridization to a *v-erbA* sequence; its function was completely unknown¹⁹. The gene is localized on chromosome 5 and expressed in a wide variety of tissues. By amino-acid sequence comparison, we found that it also resembles hERR-1 and hERR-2, two other 'orphan receptors' obtained by cross-screening and which are currently in search of a function and a ligand¹³. With this new example it may be easier to uncover the function and recognition sequence of several other receptor-related genes that have been cloned using sequence homology screening of cDNA libraries.

Finally, it is important to note that COUP-TF has been demonstrated to function *in vitro* under cell-free conditions^{2,7,8}. A second 'activator protein', S300-II, was required for this function⁸. S300-II has been purified and shown to interact functionally with COUP-TF, but S300-II does not bind DNA specifically. We conclude that COUP-TF is a member of the steroid-receptor superfamily of gene regulators that can function as a transcription factor *in vitro*^{2,7,8} □

Received 22 March; accepted 15 May 1989.

1. Sagami, I., Tsai, S. Y., Wang, H., Tsai, M.-J. & O'Malley, B. W. *Molec. cell. Biol.* **6**, 4259-4267 (1986).
2. Wang, L.-H., Tsai, S. Y., Sagami, I., Tsai, M.-J. & O'Malley, B. W. *J. biol. Chem.* **262**, 16080-16086 (1987).
3. Bagchi, M. K., Tsai, S. Y., Tsai, M.-J. & O'Malley, B. W. *Molec. cell. Biol.* **7**, 4151-4158 (1987).
4. Hwang, Y.-P., Crowe, D. T., Wang, L.-H., Tsai, S.-Y. & Tsai, M.-J. *Molec. cell. Biol.* **8**, 2070-2077 (1988).
5. Hwang, Y.-P., Wang, L.-H., Tsai, S. Y. & Tsai, M.-J. *J. biol. Chem.* **263**, 13470-13474 (1988).
6. Evans, R. M. *Science* **240**, 889-895 (1988).
7. Pastorio, M. et al. *Molec. Cell. Biol.* **6**, 2784-2791 (1986).
8. Tsai, S. Y., Sagami, I., Wang, H., Tsai, M.-J. & O'Malley, B. W. *Cell* **50**, 701-709 (1987).
9. Young, R. A. and Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194-1198 (1983).
10. Vinson, C. R., LaMarco, K. L., Johnson, P. F., Lanschulz, W. H. & McKnight, S. L. *Genes Dev.* **2**, 801-806 (1988).
11. Lawrence, C. B. & Goldman, D. A. *Computer appl. Biosci.* **4**, 25-33 (1988).
12. Hollenberg, S. M. & Evans, R. M. *Cell* **55**, 899-906 (1988).
13. Giguere, V., Yang, N., Segui, P. & Evans, R. M. *Nature* **331**, 91-94 (1988).
14. Tsai, S. Y. et al. *Cell* **55**, 361-369 (1988).
15. Denis, M., Poellinger, L., Wikstrom, A.-C. & Gustafsson, J.-A. *Nature* **333**, 686-688 (1988).
16. Kumar, V. & Chambon, P. *Cell* **55**, 145-156 (1988).
17. Sakai, D. D. et al. *Genes Dev.* **2**, 1144-1154 (1988).
18. Courey, A. J. & Tjian, R. *Cell* **55**, 887-898 (1988).
19. Miyajima, N. et al. *Nucleic Acids Res.* **16**, 11057-11074 (1988).
20. Pluskal, M. G., Prezekop, M. B. & Kavonjian, M. R. *Biotechniques* **4**, 272-283 (1986).
21. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
22. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
23. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-550 (1980).
24. Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. *Nature* **330**, 624-629 (1987).
25. Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. *Nature* **330**, 444-450 (1987).
26. de The, H., Marchio, A., Tiollais, P. & Dejean, A. *Nature* **330**, 667-670 (1987).
27. Weinberger, C. et al. *Nature* **324**, 641-646 (1986).
28. Walter, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7889-7893 (1985).
29. Baker, A. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3294-3298 (1988).
30. Arriza, J. L. et al. *Science* **237**, 268-275 (1987).
31. Hollenberg, S. M. et al. *Nature* **318**, 635-641 (1985).
32. Chang, C., Kokontis, J. & Liao, S. *Science* **240**, 324-326 (1988).
33. Lubahn, D. B. et al. *Science* **240**, 327-330 (1988).
34. Misrahi, M. et al. *Biochem. biophys. Res. Commun.* **143**, 740-748 (1987).

ACKNOWLEDGEMENTS. We thank Jennifer Liao, Drs Bonnie Dunbar, Christina Chang, and Craig Wilkinson for technical help and Dr Heng Wang for the methylation interference.

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (_____) _____

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Stable cell membrane labelling

P.K. Horan and S.E. Slezak

Binding fluorescent or radioactive reporter molecules to the lipid bilayer of cell membranes allows cell growth and trafficking to be monitored *in vivo*.

THE availability of lymphokines and other cytokines for therapeutic applications has prompted researchers to begin questioning the type of cellular interactions that occur *in vivo* and the role of specific cells in the pathogenesis or cure of a disease. Answering these questions requires knowing where cells migrate within the body and what signals "call" cells to specific sites. However, these types of studies have been limited by the lack of cell markers that are non-toxic and have sufficient stability in the cell to permit long-term tracking.

Zynaxis has developed a new technology that permits stable, reproducible cell labelling through the incorporation of highly aliphatic reporter molecules containing fluorochrome or radioisotope head groups into the lipid bilayer of cytoplasmic membranes. The probes are trapped once incorporated into the membrane, because of their inherent insolubility in aqueous environments. The peak excitation of the PKH2 fluorescent probe is 488 nm and the peak emission is 503 nm. Figure 1 shows that the relative intensity of PKH2-labelled cells is approximately 1,000 times greater than a non-labelled control sample.

Previous methods for studying cell trafficking have made use of ^{51}Cr , fluorescein isothiocyanate, rhodamine isothiocyanate, carboxy-fluorescein diacetate and Di-I-C18-(3) (refs 1, 2). Radioactive tracers such as chromium have been employed because they facilitate the localization of labelled cells without requiring microscopic examination of the tissues³. But studies using radioactive tracers have been limited because of their toxicity to some cell types and the risk of elution of the radioactive probe. Rhodamine isothiocyanate has also been used by many investigators to track cells⁴, but it elutes rapidly from cells, limiting the time they can be followed to 24–48 hours after labelling.

Stable labels

The Zynaxis fluorescent cell linkers have been used to label neurons, glial cells, bacteria, yeast, lymphocytes, monocytes, red blood cells, platelets, endothelial cells, and many tissue-cultured cells of tumour or normal cell origin. One of the major benefits of the technology is its ability to retain the probe in the cell. In cultures of mixtures of labelled and non-labelled cells in complete media, we found

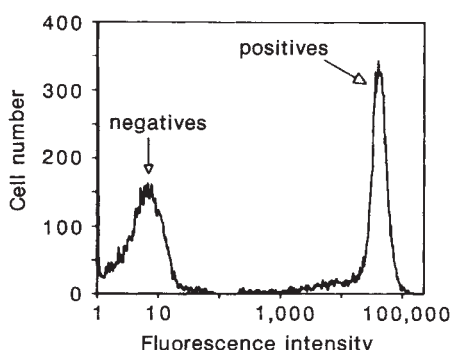


FIG. 1 Moloney leukemia virus-induced mouse lymphoma (Yac-1) cells labelled with the PKH2 fluorescent probe compared to unlabelled cells.

little or no transfer of fluorescent dye from the labelled cells to the non-labelled cells after several days in culture. James Jacobberger of Case Western Reserve University in Cleveland, Ohio has been applying the technology to growth studies of rat olfactory bulb neurons. In co-cultures of neurons and glial cells, Jacobberger has observed that labelled neural cells retain the probe, with growing axonal processes becoming brightly labelled, but no transfer of label to the glial cells takes place (personal communication).

The fluorescent label is partitioned between daughter cells as a parent cell undergoes division, allowing it to be used to identify cells of interest from an admixture of cells and to measure the growth rate between initiation and conclusion of a particular assay⁵. The fluorescence intensity of the daughter cells is approximately one-half that of the labelled parent cell. The tracking compound does not alter the doubling time of cells when used at appropriate concentrations⁵.

In vivo tracking

The fluorescent compound has also been used to label either the effector or the target cells in natural-killer cell-mediated cytotoxicity assays. The release of ^{51}Cr in these assays is unaffected, suggesting that target cell recognition and killing is not altered by labelling either cell type⁶. Michael Lotze of the US National Cancer Institute (NCI), in collaboration with Jim Yang and Steven Rosenberg, has also demonstrated that labelling mouse tumour infiltrating lymphocytes (TILs) with the fluorescent probe does not affect their ability to reject a tumour (personal communication).

Although ^{51}Cr has been used widely in

the study of circulating red blood cells, elution of the radioisotope and the resulting radiotoxicity have made this approach unpopular⁷. Using the cell linker technology, it has been possible to label stably red blood cells with a non-radioactive probe and monitor their time in circulation through fluorescence detection⁸.

John Hay and Gary Teare from the University of Toronto, Canada, have used the fluorescent cell linker technology to label lymphocytes which were obtained from sheep intestinal lymph. The lymphocytes obtained were reintroduced into the venous systems of the sheep, and the percentage and intensity of the labelled cells was monitored over a six-week period through periodic venipuncture and lymphatic drain. Fluorescence intensity analysis by flow cytometry demonstrated that the dye has a mean half-life in lymphatic cells of approximately 12 days *in vivo*.

In another application, Robert Wiltout, Anne Pilaro and Thomas Sayers of the NCI have been using the technology to study polymorphonuclear leukocyte trafficking. They have found that labelling does not interfere with chemotaxis or the abilities of the cells to respond to various stimuli *in vitro* (personal communication). They are now studying the migrational patterns of labelled cells in response to different cytokines *in vivo*.

On the horizon

The linker technology described has been applied widely to mammalian cells, but preliminary evidence suggests that it may also be useful in the study of other cell types, such as bacteria, yeast and fungi. Current applications focus primarily on *in vitro* labelling, and the subsequent study of labelled cells both *in vitro* and *in vivo*. *In vivo* labelling applications have so far been restricted to phagocytic cells^{9–11}, but the technique may be expanded to include non-phagocytic cells.

Zynaxis is now developing a molecule that can be excited at 488 nm, and that emits fluorescent light in the red region of the spectrum. This red tracking dye will be able to be used in combination with other probes such as fluorescein, and will be available at the beginning of next year. Zynaxis is also planning to introduce a ^{125}I tracking dye to be used with the fluorescent tracking molecules. The joint application of fluorescent and radioactive tracing will allow gamma-scintigraphy to

be used to localize injected cells, and the subsequent study of sectioned tissues using fluorescence image cytometry. □

Paul Karl Horan and Sue E. Slezak are at Zynaxis Cell Science, Inc., 371 Phoenixville Pike, Malvern, Pennsylvania 19355, USA. For more information, fill in reader service number 100.

1. Olszewski, W.L. *In Vivo Migration of Immune Cells*, 223-253 (CRC Press, Boca Raton, Florida, 1987).
2. Honing, M.G. & Hume, R.I. *J. Cell Biol.* **103**, 171-187 (1986).
3. Wiltroth, R.H. *et al. Cancer Immunol. Immunotherapy* **14**, 172 (1983).
4. Butcher, E.C., Scollay, R.G. & Weissman, I.L. *J. Immunol. Meth.* **37**, 109 (1980).
5. Horan, P.K., Slezak, S.E. & Jensen, B.D. (in preparation).
6. Slezak, S.E. & Horan, P.K. *J. Immunol. Meth.* **117**, 205 (1989).
7. Wintrobe, M.M. *et al. in Clinical Hematology* 8th edn (eds Lee, R.G., Bithell, T.C. & Foerster, J.) 75 (Lea & Febiger, Philadelphia, 1981).
8. Slezak, S.E. & Horan, P.K. *Blood* (in the press).
9. Melnicoff, M.J., Morahan, P.S., Jensen, B.D., Breslin, E.W. & Horan, P.K. *J. Leukemia Biol.* **43**, 387 (1988).
10. Melnicoff, M.J., Horan, P.K., Breslin, E.W. & Morahan, P.S. *J. Leukemia Biol.* **44**, 367 (1988).
11. Melnicoff, M.J., Horan, P.K. & Morahan, P.S. *Cell Immunol.* **118**, 178 (1989).

ADVERTISEMENTS

Custom DNA

Purified and Delivered in 48 hours

100 ug, \$10.00 a base for the first 15 bases, \$7.50 for each additional base. Milligram quantities available.

Research Genetics

2130 S. Memorial Parkway Huntsville, AL 35801

1-800-533-4363

In the United Kingdom 0-800-89-1393

In Canada 1-800-833-8348

FAX 205-536-9016 NW2971/D

**FOR
EXPERT
CARE & FEEDING
OF
YOUR
ANTIBODIES
CALL
800-824-8540**

**CUSTOM SERVICES
HELPING YOU OUTSMART THE HERD
SINCE 1961**

Antibodies Incorporated
DAVIS, CALIFORNIA

Copies of articles from
this publication are now
available from the UMI
Article Clearinghouse.

**UMI Article
Clearinghouse**

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

What's in a label

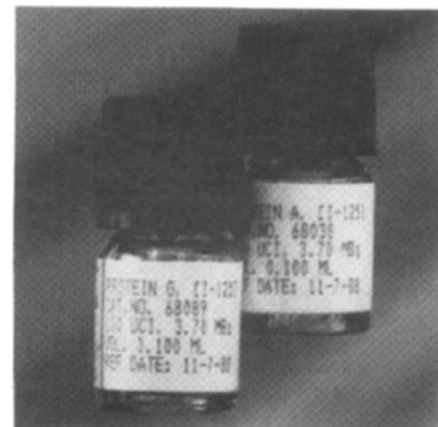
A chemiluminescent detection kit for Western blots, ¹²⁵I-labelled proteins A and G, and DAB enzyme in tablet form are just a few of this week's products for labelling applications.

EPICENTRE Technologies has announced the availability of HK phosphatase for use in DNA labelling at 5' protruding ends (Reader Service No. 101). HK phosphatase replaces calf intestinal and *Escherichia coli* bacterial alkaline phosphatases without the inherent problems of these enzymes, such as exonuclease activity. The enzyme is derived from an Antarctic bacterium which is heat-killed at 65 °C for 15 minutes. Using the reagent eliminates the need for phenol and chloroform extractions or ethanol precipitations that cause DNA loss, says Epicentre. All steps, ranging from restriction endonuclease digestion and dephosphorylation of DNA to ligation or kinasing, can be performed in one tube. HK phosphatase costs \$55 (US) for 15 units and \$240 (US) for 100 units.

TAGO, Inc. has two new **negative controls** for alkaline phosphatase and horseradish peroxidase testing procedures (Reader Service No. 102). The company says its TAGO-ZYME enzyme controls are designed to be used as non-reactive control reagents in any immunoassay that employs alkaline phosphatase and horseradish peroxidase conjugated antibody as the reporter, or signal-generating, reagent. The controls provide for both intact molecules and F(ab')₂ fragments, and typical applications include enzyme immunohistochemistry, Western or other blotting techniques, and ELISAs. TAGO recommends the use of its TAGO-ZYME reagents whenever it is necessary to rule out the possibility of non-specific reactions of the enzyme-conjugated immunoglobulins in the assay system. A 1-ml vial of either of the negative controls costs \$70 (US).

Neurons, neuronal processes, peripheral nerves, sympathetic ganglion cells and adrenal medullae are stained by Dako Corporation's new **anti-neurofilament protein antibody**, Dako-NF (Reader Service No. 103). Dako-NF recognizes the 200-kD and 70-kD components of the three major polypeptide subunits generally present in neurofilaments, without reacting with other filament proteins, says Dako. The antibody exhibits species cross-reactivity, and can be used on routinely fixed paraffin-embedded sections as well as on cryostat sections and centrifuge preparations. Dako sells 1-ml vials of Dako-NF for \$126 (US).

¹²⁵I-labelled **proteins A and G** are now available through ICN Biomedicals, for research use (Reader Service No. 104).



ICN Biomedicals' proteins A and G now come labelled with ¹²⁵I.

The company says its ¹²⁵I protein A and ¹²⁵I protein G exhibit exceptional sensitivity in the detection of antibody binding when screening lambda gt11 libraries. ICN Biomedicals sells 100 µCi of the radio-labelled protein A for \$165 (US) and 50 µCi of the protein G for \$220 (US).

Staining techniques

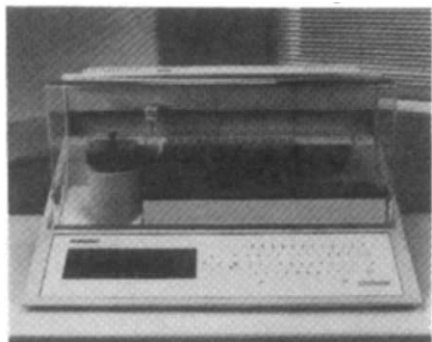
Sigma Chemical has announced the first product in its new line of **tableted enzyme substrates**, 3,3'-diaminobenzidine tetrahydrochloride (DAB), for use in immunoblotting and immunohistological staining procedures (Reader Service No. 105). Tableted enzyme substrates offer the advantages of convenience and accuracy of dosage, says Sigma. DAB is a horseradish peroxidase substrate that produces a brown insoluble end product that can be observed visually. The 10-mg enzyme substrate tablets are packaged individually in



Just drop Sigma's DAB tablets in water for horseradish peroxidase studies.

foil packets of 50 or 100, for \$71.70 (US) and \$119.55 (US), respectively. The foil packets provide enhanced stability as well as ease of use and storage, says the company. Bulk quantities are also available.

Shandon says its Cadenza **immunostainer** can be used to replace manual immunostaining methods (*Reader Service No. 106*). The instrument is simple to load and can carry out complex and lengthy procedures involving reagent additions, incubations and washings, while protecting the specimen, says the company. The Cadenza can process up to 20 individual samples per run, giving an average daily throughput of 40 slides. It holds up to 18 separate reagents in addition to a buffer wash. Special well-shaped disposable coverplates fit over the specimen slides to distribute 100 μ l of staining reagent. As a new solution is introduced, it displaces the



Shandon's new Cadenza immunostainer can process up to 40 slides per day.

previous fluid into a waste tray. The \$25,000 (US) Cadenza can also be programmed for a 50 μ l incubation step.

Four new **immunohistostaining systems** available in kit form from Dynatech incorporate a nuclear counterstain that allows the visualization of enzyme activity and cell morphology (*Reader Service No. 107*). The company offers red and blue reagents for use with alkaline phosphatase-labelled antibodies, and an orange and black detection system for horseradish peroxidase conjugates. Dynatech says the reagent formulations provide strong image contrast with a low background uptake, making them suitable for frozen and paraffin-embedded tissues, as well as other cellular preparations. Each kit contains sufficient chromagen solution, activator solution, buffered substrate solution and counterstain for 1,000 slides, plus a detailed protocol for localizing peroxidase or alkaline phosphatase reporter systems.

The light slide

Perkin-Elmer says its new model LS-40 programmable **fluorescence LC detector** can detect fluorescence down to the femto-gram level, and also has the capability to detect phosphorescence, chemiluminescence and bioluminescence (*Reader Service No. 108*). The unit features a 4- μ l

illuminated flow cell and a compact dual monochromator system. Excitation and emission monochromators can be time-programmed for compound-specific detection. The \$12,000 (US) model LS-40 also features a pre-scan function that automatically determines the maximum excitation and emission wavelengths.

Southern-Light is the name of a new **chemiluminescent detection kit** from Tropix, Inc. for use with membrane-based DNA, such as Southern blots (*Reader Service No. 109*). It can be used with a variety of DNA probes, including both biotinylated and directly conjugated alkaline phosphatase probes. Each test kit is supplied with material sufficient to perform 50 10 $\times$ 10 cm² blots, including AMPPD, a chemiluminescent enzyme substrate designed for use with alkaline phosphatase, and an optimized blocking buffer system. Tropix says results obtained using the \$175 (US) Southern-Light kit can be imaged on instant or X-ray film; hard copy results are obtainable following room temperature incubation within 5–60 minutes.

Cambridge Instruments has a new **vertical fluorescence illuminator** for its DiaStar compound microscope that can be used for both incident and transmitted light fluorescence microscopy (*Reader Service No. 110*). Rapid transition from

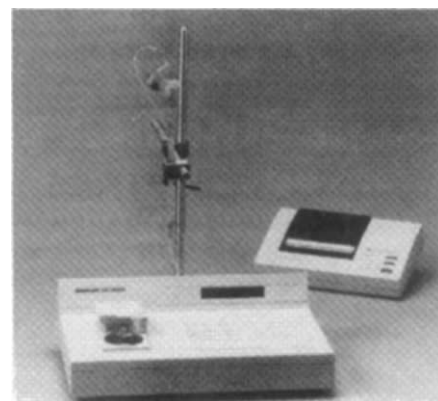


Vertical fluorescence illumination is new to Cambridge Instruments' DiaStar microscope.

transmitted light bright field to contrast applications is also possible. The attachment features a high-performance FITC fluor cluster consisting of an exciter filter, a barrier filter and a dichroic mirror pre-assembled into a fluorescence cube. The complete vertical illuminator outfit consists of a fluor cluster, lamp holder, power supply, lamp, low fluorescence immersion oil and a centration target to help align the fluorescence illuminator.

Radiochemistry

Bioscan has an **LC radioisotope flow analysis module** for use with its Quick-Count benchtop radioisotope counters (*Reader Service No. 111*). The model FL-2000 module provides a compact hardware and software system with microlitre



Bioscan now offers an LC radioisotope flow analysis module with its QuickCount counter.

to millilitre sample volume. The flow cell is configured using readily available standard tubings. Bioscan says the software provides easy setup, data tabulations, peak integrations and scaled plots, yet retains the flexibility to count single samples. ³²P and ¹²⁵I are counted directly with the unit without the use of scintillation fluid or crystals. The flow analysis module sells for \$18,000 (US).

ISS-Enprotech is expanding its line of **electrophoresis-grade radiochemicals** to include methionine L-³⁵S and cysteine L-³⁵S (*Reader Service No. 112*). Both radio-labelled amino acids have a specific activity of 600 Ci mmol⁻¹. Methionine L-³⁵S is packaged in 50 mM tricine (pH 7.4) 10 mM 2-mercaptoethanol, and cysteine L-³⁵S is packaged in 10 mM dithiothreitol

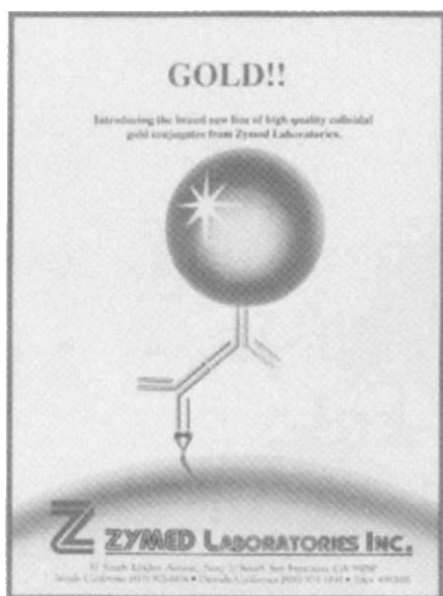


ISS-Enprotech has electrophoresis-grade methionine L-³⁵S and cysteine L-³⁵S.

in water. The labelled methionine costs \$350 (US) for 5 mCi; labelled cysteine costs \$395 (US) for 5 mCi.

As good as gold

Zymed is introducing a new line of **colloidal gold conjugates** for electron microscopy, light microscopy and blotting applications (*Reader Service No. 113*). The line includes gold conjugates of protein A, protein G, streptavidin and antibodies to rabbit IgG, mouse IgG, mouse IgM, mouse IgG plus IgM, goat IgG,



The book on Zymed's colloidal and particle gold products.

guinea pig IgG and rat IgG. Particle sizes of 1, 5, 10, 15, 20 and 30 nm are available. Prices range from \$28 (US) for blotting grade conjugates and \$46 (US) for light microscopy particles, to \$65 (US) for gold conjugates suitable for electron microscopy.

Biocell is expanding its range of **immunogold reagents** to include gold-conjugated antibodies to mouse IgG plus IgM, human IgG (gamma), guinea pig IgG, sheep IgG, goat IgG and protein G (*Reader Service No. 114*). The company is also introducing **cationic colloidal gold** sols with particle sizes of 2, 5, 10, 15, 20, 40 and 80 nm, for use in the one-step detection of subcellular anionic sites in cells and tissues. Biocell says its antibody-gold conjugates can detect less than 1 pg of protein in immunoblotting tests, with minimal cross-reactivity. Conjugates for electron microscopy are available with gold particle sizes of 5, 10, 15 and 20 nm; Biocell says the preparations are at least 85 per cent singlets. Prices for the antibody conjugates range from £50 (UK) to £120 (UK) for 1 ml.

Sera-Lab Ltd is now offering its **colloidal gold conjugates** in 250-µl vial sizes to accommodate smaller laboratories (*Reader Service No. 115*). The company says the new size is ideal for those laboratories which find that its highly concentrated 1-ml size offers too many tests.

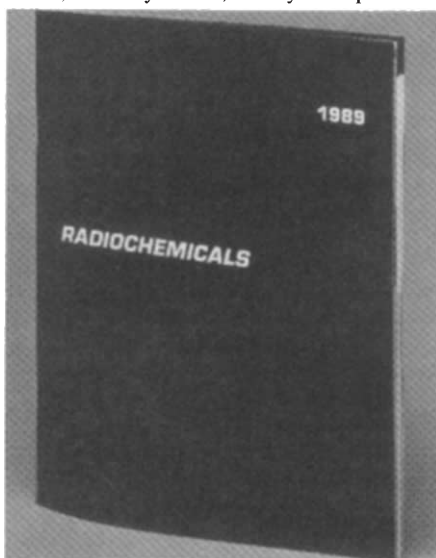
Mail-order means

A new catalogue featuring nearly 400 **immunological reagents**, lectins and biotin/avidin system products is now available from Vector Laboratories (*Reader Service No. 116*). New listings include streptavidin conjugates, affinity-purified secondary antibodies labelled with fluorochromes

or enzymes, a more sensitive Vectastain ABC series, new conjugated and unconjugated lectins, biotinylated DNA and protein molecular weight markers, anti-avidin and anti-biotin products, and biotin labelling reagents for DNA, RNA, proteins or carbohydrates. Labels include peroxidase, alkaline phosphatase, glucose oxidase, fluorescein, rhodamine, Texas Red, phycoerythrin and a new blue fluorescent label referred to as AMCA.

The "May We Help You?" literature series from MSD Isotopes offers application notes and helpful hints for the use of the company's line of **stable isotopes** (*Reader Service No. 117*). Topics include the care and handling of NMR solvents, an NMR solvent table, using deuterated compounds as FAB solvents, preparing diacetone glucose for carbohydrate MS, using labelled pollutants as isotope dilution standards, isotope labelled spin label compounds for ESR, and details on specific labelled compounds, such as ^{13}C , ^{15}N , and ^{17}O and ^{18}O .

Sigma Chemical's 1989 **radiochemicals catalogue** contains an expanded list of biochemicals, including ^{14}C -labelled amino acids, carbohydrates, methylated proteins



Sigma's 1989 catalogue on radiochemicals, and molecular weight markers, and ^3H -labelled nucleotides (*Reader Service No. 118*). A wide range of photographic equipment, supplies for autoradiography, and scintillation vials and caps are among the new products listed. Sigma also offers a custom synthesis service, and a catalogue detailing over 300 products that were previously only available on a custom basis.

The Atlantic Antibodies division of INCSTAR has just released its free 1989 catalogue of **immunoassay reagents** (*Reader Service No. 119*). The company offers products for both diagnostic and research applications, including IgG frac-



INCSTAR's complete catalogue.

tions, fluorescent conjugates and enzyme conjugates. Atlantic Antibodies also provides technical services and custom immunization. □

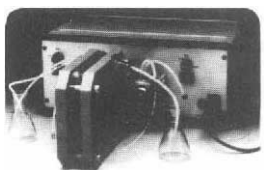
These notes are compiled by Cary Prince from information provided by the manufacturers. To obtain additional information about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID

NEW



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.45

Economical Modular Tissue Culture Incubator



- Safe for AIDS research • Reliable, for In-Vitro fertilization
- Versatile enough for thousands of lab uses

Easy to use — Just flush with gas mixture, seal and place at approp. temp. Perform multiple experiments simultaneously w/out cross contamination. Thousands in use worldwide.

For information contact: **Flow Laboratories, Ltd.** or, **Billups-Rothenberg, Inc.** • P.O. Box 977
Del Mar, CA 92014-0977 USA • (619) 755-3309

Reader Service No.37